

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number
WO 02/072779 A2

- (51) International Patent Classification⁷: **C12N**
- (21) International Application Number: **PCT/US02/07584**
- (22) International Filing Date: **12 March 2002 (12.03.2002)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
60/275,202 **12 March 2001 (12.03.2001)** **US**
- (71) Applicant (for all designated States except US):
AFFYMETRIX, INC. [US/US]; 3380 Central Expressway, Santa Clara, CA 95051 (US).
- (71) Applicants and
- (72) Inventors: **MCGALL, Glenn** [US/US]; 1121 Sladky

Avenue, Mountain View, CA 94043 (US). **BARONE, Anthony, D.** [US/US]; 2118 Ellen Avenue, San Jose, CA 95125 (US).

(74) Agent: **VIKSINS, Ann, S.**; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).

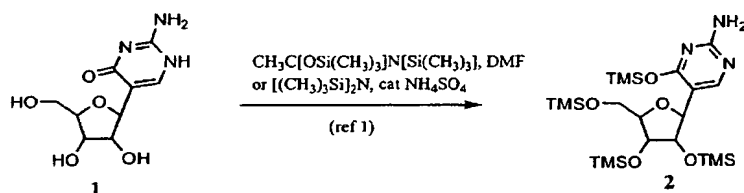
(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

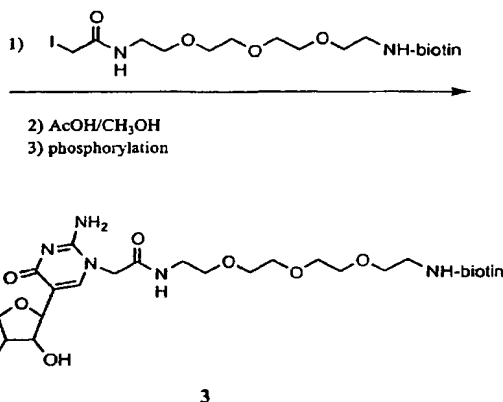
[Continued on next page]

(54) Title: **NUCLEIC ACID LABELING COMPOUNDS**

fluorescein and biotin labeled N1-ψCTP



(57) Abstract: Nucleic acid labeling compounds are disclosed. The compounds are synthesized by condensing a heterocyclic derivative with a cyclic group (e.g. a ribofuranose derivative). The labeling compounds are suitable for enzymatic attachment to a nucleic acid, either terminally or internally, to provide a mechanism of nucleic acid detection.





European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *without international search report and to be republished upon receipt of that report*

NUCLEIC ACID LABELING COMPOUNDS

5

Statement of Rights to Inventions made under Federally Sponsored Research

This invention was made with Government support under contract 70NANB5H1031 awarded by the Advanced Technology Program of the
10 National Institute of Standards and Technology. The Government has certain rights in this invention.

CROSS REFERENCE TO RELATED APPLICATIONS

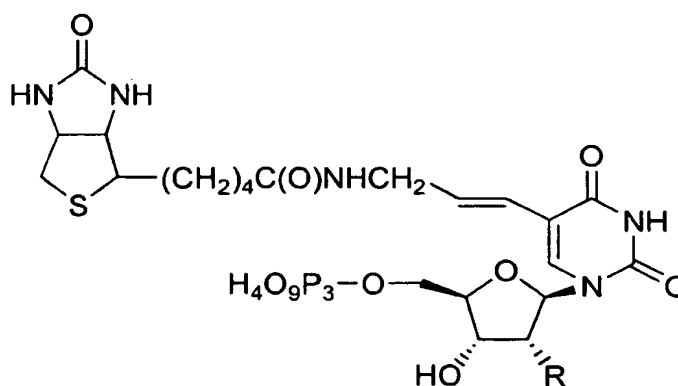
This application claims priority from U.S. provisional patent Application Serial No. 60/275,202, filed on March 12, 2001, which is herein
15 incorporated by reference for all purposes.

BACKGROUND OF THE INVENTION

Gene expression in diseased and healthy individuals is oftentimes different and characterizable. The ability to monitor gene expression in such cases provides medical professionals with a powerful diagnostic tool. This form
20 of diagnosis is especially important in the area of oncology, where it is thought that the overexpression of an oncogene, or the underexpression of a tumor suppressor gene, results in tumorigenesis. See Mikkelsen et al. *J. Cell. Biochem.* 1991, 46, 3-8.

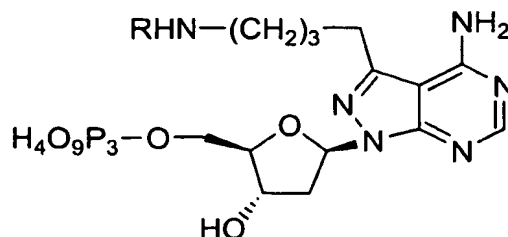
One can indirectly monitor gene expression, for example, by
25 measuring a nucleic acid (e.g., mRNA) that is the transcription product of a targeted gene. The nucleic acid is chemically or biochemically labeled with a detectable moiety and allowed to hybridize with a localized nucleic acid probe of known sequence. The detection of a labeled nucleic acid at the probe position indicates that the targeted gene has been expressed. See, e.g., International
30 Application Publication Nos. WO 97/27317, WO 92/10588 and WO 97/10365.

The labeling of a nucleic acid is typically performed by covalently attaching a detectable group (label) to either an internal or terminal position. Scientists have reported a number of detectable nucleotide analogues that have been enzymatically incorporated into an oligo- or polynucleotide. Langer et al.,
 5 for example, disclosed analogues of dUTP and UTP that contain a covalently bound biotin moiety. *Proc. Natl. Acad. Sci. USA* **1981**, 78, 6633-6637. The analogues, shown below, possess an allylamine linker arm that is attached to the C-5 position of the pyrimidine ring. The dUTP and UTP analogues, wherein R is H or OH, were incorporated into a polynucleotide.



10

Petrie et al. disclosed a dATP analogue, 3-[5-[(N-biotinyl-6-aminocaproyl)-amino]pentyl]-1-(2-deoxy-β-D-erythro-pentofuranosyl)-1H-pyrazolo[3,4-d]pyrimidin-4-amine-5'-triphosphate. *Bioconjugate Chem.* **1991**,
 15 2, 441-446. The analogue, shown below, is modified at the 3-position with a linker arm that is attached to a biotin moiety. Petrie et al. reported that the compound wherein R is biotin is incorporated into DNA by nick translation.

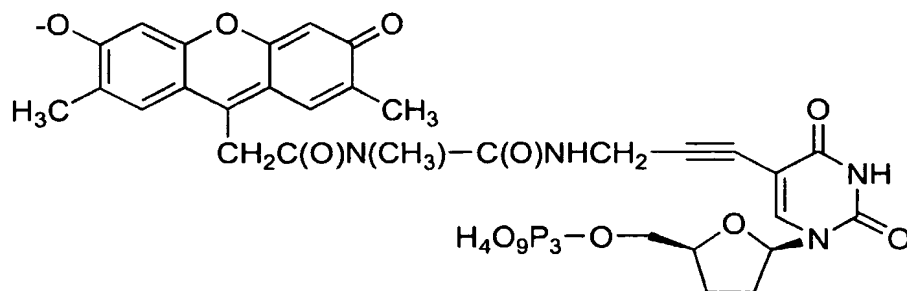


20

Prober et al. disclosed a set of four dideoxynucleotides, each containing a succinylfluorescein dye. *Science* **1987**, 238, 336-341. The

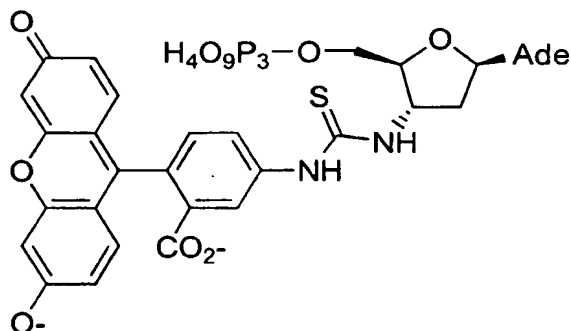
dideoxynucleotides, one of which is shown below, were enzymatically incorporated into an oligonucleotide through a template directed extension of a primer. The compounds provided for a DNA sequencing method based on gel migration.

5



Herrlein et al. disclosed modified nucleoside triphosphates of the four DNA bases. *Helv. Chim. Acta* **1994**, 77, 586-596. The compounds, one of which is shown below, contain a 3'-amino group containing radioactive or fluorescent moieties. Herrlein et al. further described the use of the nucleoside analogues as DNA chain terminators.

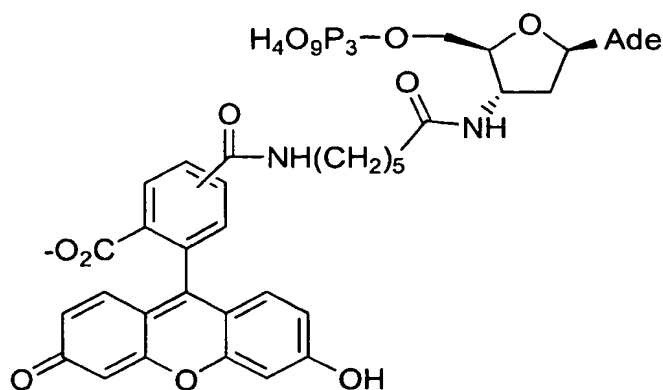
10



15

Cech et al. disclosed 3'-amino-functionalized nucleoside triphosphates. *Collect. Czech. Chem. Commun.* **1996**, 61, S297-S300. The compounds, one of which is shown below, contain a fluorescein attached to the 3'-position through an amino linker. Cech et al. proposed that the described functionalized nucleosides would be useful as terminators for DNA sequencing.

20



The development of novel nucleic acid labeling compounds that are effectively incorporated into a nucleic acid to provide a readily detectable composition would benefit genetic analysis technologies. It would aid, for example, in the monitoring of gene expression and the detection and screening of mutations and polymorphisms. Such a compound should be suitable for enzymatic incorporation into a nucleic acid. Furthermore, the nucleic acid to which the labeling compound is attached should maintain its ability to bind to a probe, such as a complementary nucleic acid.

Although nucleic acid labeling compounds for use as coupling agents for probes are available there is a continuing need for additional compounds that are more efficient labeling compounds. There also exists a need for compounds that have increased solubility. This will make the compounds more useful for monitoring gene expression.

SUMMARY OF THE INVENTION

The present invention relates to nucleic acid labeling compounds. More specifically, the invention provides compounds containing a detectable moiety. The invention also provides methods of making these compounds. It further provides methods of attaching the compounds to a nucleic acid. The nucleic acid labeling compounds or the present invention are effectively incorporated into a nucleic acids to provide readily detectable compositions that are useful for genetic analysis technologies. These compounds and the detectable compositions can aid, for example, in the monitoring of gene

expression and the detection and screening of mutations and polymorphisms. Thus, the compounds of the invention are suitable for enzymatic incorporation into nucleic acids. Furthermore, the nucleic acids to which the labeling compound are attached maintain their ability to bind to a probe, such as, for
 5 example a complementary nucleic acid.

The present invention provides nucleic acid labeling compounds that are capable of being enzymatically incorporated into a nucleic acid. The nucleic acids to which the compounds are attached maintain their ability to bind to a complementary nucleic acid sequence.

10

BRIEF DESCRIPTION OF THE DRAWINGS

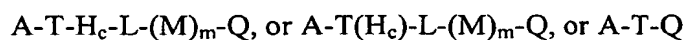
Figure 1 illustrates a non-limiting set of template moieties.

Figure 2-10 illustrate synthetic routes to fluorescein and biotin labeled compounds of the invention.

DETAILED DESCRIPTION OF THE INVENTION

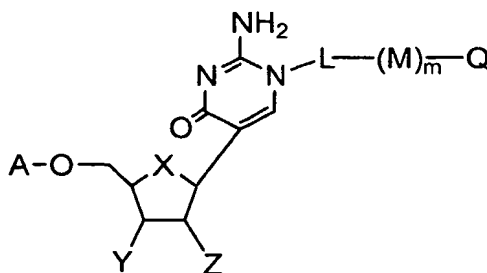
15

The nucleic acid labeling compounds of the present invention have the following structure:

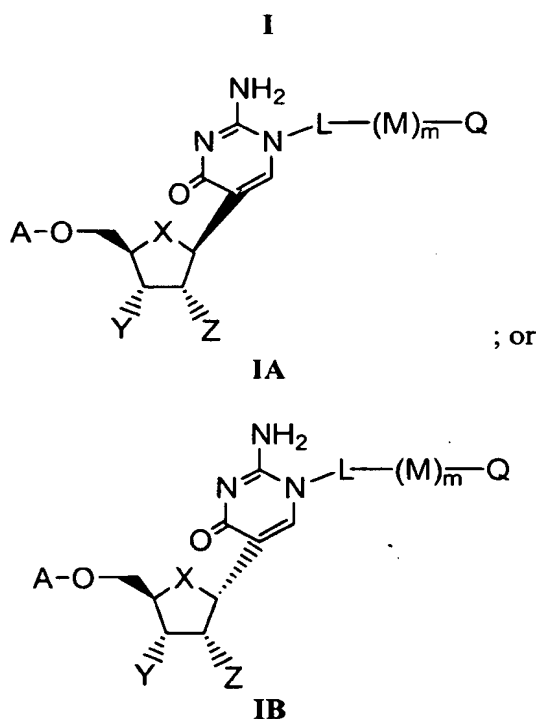


wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid; T is a template moiety; H_c is a
 20 heterocyclic group; L is a linker moiety; each M is a connecting group; Q is a detectable moiety.

In one embodiment, the nucleic acid labeling compounds have the following structures:



25



5

wherein A is hydrogen or a functional group that permits the
 10 attachment of the nucleic acid labeling compound to a nucleic acid; X is O, S,
 NR₁ or CHR₂, wherein R₁ and R₂ are, independently, hydrogen, alkyl or aryl; Y
 is hydrogen, N₃, F, OR₉, SR₉ or NHR₉, wherein R₉ is hydrogen, alkyl or aryl; Z
 is hydrogen, N₃, F, OR₉, SR₉ or NHR₁₀, wherein R₁₀ is hydrogen, alkyl or aryl; L
 is a linker moiety; Q is a detectable moiety; and M is a connecting group,
 15 wherein m is an integer ranging from 0 to about 20.

In another embodiment, A is H or H₄O₉P₃⁻; X is O; Y is H or OR₉,
 wherein R₉ is hydrogen, alkyl or aryl; Z is hydrogen, N₃, F or OR₁₀, wherein R₁₀
 is hydrogen, alkyl or aryl; and L is amido alkyl and M is -(M₁)_a-(M₂)_b- wherein a
 and b are independently integers from 0 to about 5 and the sum of a and b is not
 20 greater than 15.

In another embodiment, L is -(CH₂)_nC(O)NR₃-, wherein R₃ is
 hydrogen, alkyl or aryl and n is an integer ranging from about 1 to about 10; M₁
 is -(CH₂)_iO- and M₂ is -(CH₂)_jNH-, wherein i and j are independently integers
 from 1 to about 5.

In another embodiment, Y is H or OH; Z is H or OH; -L is -CH₂C(O)NH-; M₁ is -(CH₂CH₂O)₃- and M₂ is -CH₂CH₂NH-; and Q is biotin or a carboxyfluorescein.

In another embodiment, Y is OH; Z is OH; -L-(M)_m- is
 5 -CH₂C(O)NH-(CH₂CH₂O)₃-CH₂CH₂NH-; and Q is biotin or Y is OH; Z is OH; -L-(M)_m- is -CH₂C(O)NH-(CH₂CH₂O)₃-CH₂CH₂NH-; Q is carboxyfluorescein.

In another embodiment, A is H or H₄O₉P₃-; X is O; Y is H or OR₉, wherein R₉ is hydrogen, alkyl or aryl; Z is hydrogen, N₃, F or OR₁₀, wherein R₁₀ is hydrogen, alkyl or aryl; and L is amido alkyl and M is
 10 -(M₁)_a-(M₂)_b-(M₃)_c-(M₄)_d wherein a, b, c, and d are independently integers from 0 to about 5 and the sum of a, b, c, and d is not greater than 15.

In another embodiment, L is -(CH₂)_nC(O)NR₃-, wherein R₃ is hydrogen, alkyl or aryl and n is an integer ranging from about 1 to about 10; each M is independently -C(O)(CH₂)_kO-, -(CH₂)_iO- or -(CH₂)_jNH-, wherein i, j
 15 and k are independently integers from 1 to about 5.

In another embodiment, L is -CH₂C(O)NR₃-, wherein R₃ is hydrogen or, alkyl; M₁ is -(CH₂)₂NH-, M₂ is -C(O)(CH₂)₂O-, M₃ is -(CH₂CH₂O)₃- and M₄ is -(CH₂)₂NH-.

In another embodiment, Y is OH; Z is OH; -L-(M)_m- is
 20 -CH₂C(O)NH-CH₂CH₂NH-C(O)CH₂CH₂O-(CH₂CH₂O)₃-CH₂CH₂NH-; Q is biotin.

In another embodiment, Y is OH; Z is OH; -L-(M)_m- is -CH₂C(O)NH-CH₂CH₂NH-; m is 1; and Q is carboxyfluorescein.

In another embodiment, L is -(CH₂)₂C(O)NR₃-, wherein R₃ is
 25 hydrogen or, alkyl; M₁ is -(CH₂)₂NH-, M₂ is -C(O)(CH₂)₂O-, M₃ is -(CH₂CH₂O)₃- and M₄ is -(CH₂)₂NH-.

In another embodiment, Y is OH; Z is OH; -L-(M)_m- is -CH₂C(O)NH-CH₂CH₂NH-C(O)CH₂CH₂O-(CH₂CH₂O)₃-CH₂CH₂NH-; Q is biotin.

In another embodiment, Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{CH}_2\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-$; m is 1; and Q is carboxyfluorescein.

In another embodiment, A is H or $\text{H}_4\text{O}_9\text{P}_3-$; X is O; Y is H or OR_9 , wherein R_9 is hydrogen, alkyl or aryl; Z is hydrogen, N_3 , F or OR_{10} , wherein R_{10} is hydrogen, alkyl or aryl; and L is functionalized alkyl and M is $-(M_1)_a-(M_2)_b-(M_3)_c-(M_2)_d-$ wherein a, b, c, and d are independently integers from 0 to about 5 and the sum of a, b, c, and d is not greater than 15.

In another embodiment, L is $-(\text{CH}_2)_n\text{O}-$, wherein n is an integer ranging from about 1 to about 10; each M is independently $-\text{C}(\text{O})(\text{CH}_2)_k\text{O}-$, $-(\text{CH}_2)_i\text{O}-$ or $-(\text{CH}_2)_j\text{NH}-$, wherein i, j and k are independently integers from 1 to about 5.

In another embodiment, L is $-\text{CH}_2\text{O}-$; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$ and M_4 is $-(\text{CH}_2)_2\text{NH}-$.

In another embodiment, Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{O}-\text{CH}_2\text{CH}_2\text{NH}-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; Q is biotin or Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{O}-\text{CH}_2\text{CH}_2\text{NH}-$; m is 1; and Q is carboxyfluorescein.

In another embodiment, L is $-(\text{CH}_2)_n\text{NR}_3-$, wherein R_3 is hydrogen, alkyl or aryl and n is an integer ranging from about 1 to about 10; each M is independently $-\text{C}(\text{O})(\text{CH}_2)_k\text{O}-$, $-(\text{CH}_2)_i\text{O}-$ or $-(\text{CH}_2)_j\text{NH}-$, wherein i, j and k are independently integers from 1 to about 5.

In another embodiment, L is $-(\text{CH}_2)_6\text{NH}-$; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, and M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$. In another embodiment, Y is OH; Z is OH; $-L-(M)_m-$ is $-(\text{CH}_2)_6\text{NH}-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; and Q is biotin or Y is OH; Z is OH; L is $-(\text{CH}_2)_6\text{NH}-$; m is 0; and Q is carboxyfluorescein.

In another embodiment, A is H or $\text{H}_4\text{O}_9\text{P}_3-$; X is O; Y is H or OR_9 , wherein R_9 is hydrogen, alkyl or aryl; Z is hydrogen, N_3 , F or OR_{10} , wherein R_{10} is hydrogen, alkyl or aryl; and L is alkenyl alkyl and M is

$-(M_1)_a-(M_2)_b-(M_3)_c-(M_4)_d-$ wherein a, b, c, and d are independently integers from 0 to about 5 and the sum of a, b, c, and d is not greater than 15.

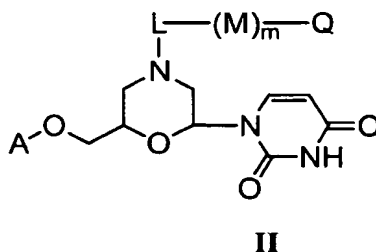
In another embodiment, L is - structure

$-\text{CH}=\text{CH}-(\text{CH}_2)_n\text{C}(\text{O})\text{N}(\text{R}_6)-$, wherein R_6 is hydrogen, alkyl or aryl and n is an integer ranging from about 0 to about 10; each M is independently $-\text{C}(\text{O})(\text{CH}_2)_k\text{O}-$, $-(\text{CH}_2)_i\text{O}-$ or $-(\text{CH}_2)_j\text{NH}-$, wherein i, j and k are independently integers from 1 to about 5.

In another embodiment, L is $-\text{CH}=\text{CH}-\text{C}(\text{O})\text{N}(\text{H})-$; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$ and M_4 is $-(\text{CH}_2)_2\text{NH}-$.

In another embodiment, Y is OH; Z is OH; $-\text{L}-(\text{M})_m-$ is $-\text{CH}=\text{CH}-\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; Q is biotin or Y is OH; Z is OH; $-\text{L}-(\text{M})_m-$ is $-\text{CH}=\text{CH}-\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-$; m is 1; and Q is carboxyfluorescein.

In another embodiment, the nucleic acid labeling compounds have the following structure:



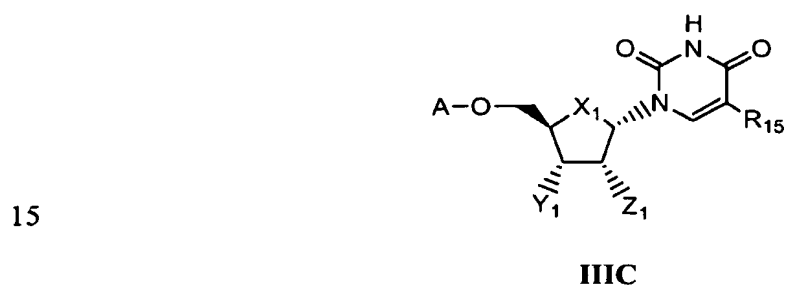
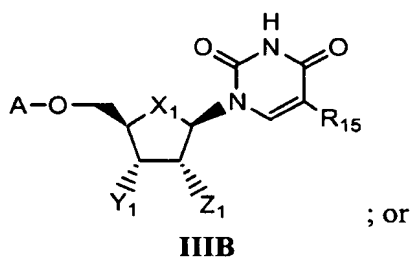
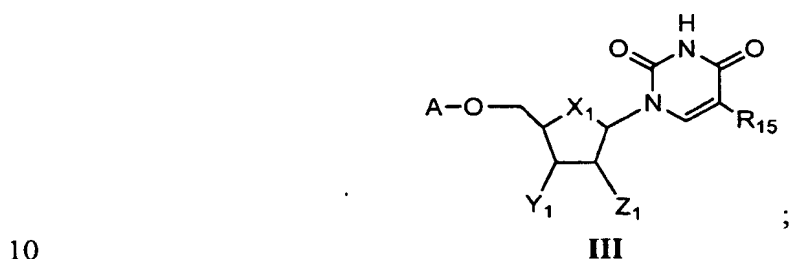
wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid; L is a linker moiety; Q is a detectable moiety; and M is a connecting group, wherein m is an integer ranging from 0 to about 3.

In another embodiment, A is H or $\text{H}_4\text{O}_9\text{P}_3-$; L is alkanoyl; and Q is biotin or a carboxyfluorescein; M is $-\text{NR}_3-$ where R_3 is hydrogen or alkyl, and wherein m is 1 or 0.

In another embodiment, L is $-\text{C}(\text{O})(\text{CH}_2)_n-$ wherein n is an integer ranging from about 1 to about 10; M is $-\text{NH}-$; Q is biotin or a carboxyfluorescein.

In another embodiment, $-\text{L}-(\text{M})_m-$ is $-\text{C}(\text{O})(\text{CH}_2)_6-\text{NH}-$; and Q is biotin or $-\text{L}-(\text{M})_m-$ is $-\text{C}(\text{O})(\text{CH}_2)_6-\text{NH}-$; and Q is carboxyfluorescein.

In another embodiment, the nucleic acid labeling compounds have the following structures:



wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid; X_1 is O, S, NR_1 or CHR_2 , wherein R_1 and R_2 are, independently, hydrogen, alkyl or aryl; R_{15} is hydrogen, alkyl or aryl; Y_1 is hydrogen, N_3 , F, OR_9 , SR_9 or NHR_9 , wherein R_9 is hydrogen, alkyl or aryl; Z_1 is hydrogen, N_3 , F, OR_9 , SR_9 or NHR_{10} , wherein R_{10} is hydrogen, alkyl or aryl; wherein one of Y_1 or Z_1 is a group having the formula $-\text{X}_2\text{L}-(\text{M})_m-\text{Q}$; X_2 is O, S, or NR_{16} , and R_{16} is hydrogen, alkyl or

aryl; L is a linker moiety; Q is a detectable moiety; and M is a connecting group, wherein m is an integer ranging from 0 to about 20.

In another embodiment, A is H or $H_4O_9P_3^-$; X_1 is O; Y_1 is hydrogen, OR_9 , wherein R_9 is hydrogen, or alkyl, or $-NR_{16}-(M)_m-Q$; wherein R_{16} is
 5 hydrogen, alkyl or aryl; Z_1 is hydrogen OR_{10} , wherein R_{10} is hydrogen, or alkyl, or $-NR_{16}-L-(M)_m-Q$, wherein R_{16} is hydrogen, or alkyl; L is N-alkyl amido; R_{15} is hydrogen or alkyl; M is $-(CH_2)_nNR_3-$ where R_3 is hydrogen or alkyl, and m is from 1 to about 15.

In another embodiment, Y_1 is hydrogen or OH; Z_1 is $-NH-L-(M)_m-Q$;
 10 L is $-(CH_2)_nNHC(O)-$ where n is an integer from 1 to about 10; M is $-NH(CH_2)_p-$ where p is an integer from 1 to about 10; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.

In another embodiment, Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.

15 In another embodiment, Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.

In another embodiment, Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.

In another embodiment, Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is
 20 $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.

In another embodiment, Y_1 is $-NH-L-(M)_m-Q$; Z_1 is hydrogen or OH; L is $-(CH_2)_nNHC(O)-$ where n is an integer from 1 to about 10; M is $-NH(CH_2)_p-$ where n is an integer from 1 to about 10; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.

25 In another embodiment, Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.

In another embodiment, of the nucleic acid labeling compound, Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.

In another embodiment, Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.

In another embodiment, Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.

- 5 In another embodiment, A is H or $H_4O_9P_3-$; X_1 is O; Y_1 is hydrogen, OR_9 , wherein R_9 is hydrogen, or alkyl, or $-O-L-(M)_m-Q$; Z_1 is hydrogen OR_{10} , wherein R_{10} is hydrogen, or alkyl, or $-O-L-(M)_m-Q$, or alkyl; L is alkylene; and m is from 1 to about 10.

- 10 In another embodiment, Y_1 is hydrogen or OH; Z_1 is $-O-L-(M)_m-Q$; L is $-(CH_2)_n-$ where n is an integer from 1 to about 12; M is $-NH-$; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.

In another embodiment, Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.

- 15 In another embodiment, wherein Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is a carboxyfluorescein.

In another embodiment, Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.

In another embodiment, Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is a carboxyfluorescein.

- 20 In another embodiment, Y_1 is $-O-L-(M)_m-Q$; Z_1 is hydrogen or OH; L is $-(CH_2)_n-$ where n is an integer from 1 to about 12; M is $-NH-$; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.

In another embodiment, Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.

- 25 In another embodiment, Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is a carboxyfluorescein.

In another embodiment, Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.

In another embodiment, Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is a carboxyfluorescein.

In another embodiment, A is H or $H_4O_9P_3-$; X_1 is O; Y_1 is hydrogen, OR₉, wherein R_9 is hydrogen, or alkyl, or $-S-L-(M)_m-Q$; Z_1 is hydrogen OR₁₀,
 5 wherein R_{10} is hydrogen, or alkyl, or $-S-L-(M)_m-Q$; L is alkylene; X_2 is S; and m is from 1 to about 10.

In another embodiment, Y_1 is hydrogen or OH; Z_1 is $-S-L-(M)_m-Q$; L is $-S-(CH_2)_n-$ where n is an integer from 1 to about 10; M is $-NH-$; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.

10 In another embodiment, Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is biotin.

In another embodiment, Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.

In another embodiment, Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is
 15 $-S-(CH_2)_2-NH-$; and Q is biotin.

In another embodiment, Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.

In another embodiment, Y_1 is $-S-L-(M)_m-Q$; Z_1 is hydrogen or OH; L is $-S-(CH_2)_n-$ where n is an integer from 1 to about 10; M is $-NH-$; R_{15} is
 20 hydrogen or methyl; and Q is biotin or a carboxyfluorescein.

In another embodiment, Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is biotin.

In another embodiment, Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.

25 In another embodiment, Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is biotin.

In another embodiment, Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.

In another embodiment, the nucleic acid labeling compounds have the following structure:



wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid; X_1 is O, S, NR_1 or CHR_2 , wherein R_1 and R_2 are, independently, hydrogen, alkyl or aryl; X_2 is a bond or alkylene; Q is a detectable moiety; and G is $-\text{L}-(\text{M})_m-$ where L is a linker moiety and each M a connecting group, where m is from 0 to about 20.

10

In another embodiment, A is H or $\text{H}_4\text{O}_9\text{P}_3-$; X_1 is O; X_2 is a bond; and G is $-\text{C}(\text{O})\text{NR}_3-(\text{CH}_2)_m-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-\text{C}(\text{O})\text{NR}_3-$, where R_3 is hydrogen or alkyl, and m and n are independently an integer from 1 to about 15.

15 In another embodiment, G is



where m is from 1 to about 6 and n is from 1 to about 4. Preferably, n is 3 or 4 and m is 5 or 6.

In another embodiment, A is H or $\text{H}_4\text{O}_9\text{P}_3-$; X_1 is O; X_2 is CH_2 ; and G is $-\text{C}(\text{O})\text{NR}_3-(\text{CH}_2)_m-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-\text{C}(\text{O})\text{NR}_3-$, where R_3 is hydrogen or alkyl, and m and n are independently an integer from 1 to about 15.

20

In another embodiment, A is H or $\text{H}_4\text{O}_9\text{P}_3-$; X_1 is O; X_2 is CH_2 ; and G is $-\text{C}(\text{O})\text{NR}_3-(\text{CH}_2)_m-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-\text{C}(\text{O})\text{NR}_3-$, where R_3 is hydrogen or alkyl, and m and n are independently an integer from 1 to about 15.

25 The present invention also provides nucleic acid derivatives produced by coupling a nucleic acid labeling compound with a nucleic acid and hybridization products comprising the nucleic acid derivatives bound to a complementary probe.

30 The present invention also provides nucleic acid derivatives produced by coupling one of the nucleic acid labeling compounds of the invention with a

nucleic acid and the hybridization products comprising the nucleic acid derivatives bound to a complementary probe. The hybridization product formed from the nucleic acid derivatives comprise the nucleic acid derivative bound to a complementary probe. In one embodiment, the probe is attached to a glass chip.

5 The present invention also provides methods of synthesizing nucleic acid derivatives by attaching a nucleic acid labeling compound to a nucleic acid. It further provides methods of detecting nucleic acids involving incubating the nucleic acid derivatives with a probe.

10 In yet another embodiment, the methods involve the steps of: (a) providing at least one nucleic acid coupled to a support; (b) providing a labeled moiety capable of being coupled with a terminal transferase to said nucleic acid; (c) providing said terminal transferase; and (d) coupling said labeled moiety to said nucleic acid using said terminal transferase.

15 In still another embodiment, the methods involve the steps of: (a) providing at least two nucleic acids coupled to a support; (b) increasing the number of monomer units of said nucleic acids to form a common nucleic acid tail on said at least two nucleic acids; (c) providing a labeled moiety capable of recognizing said common nucleic acid tails; and (d) contacting said common nucleic acid tails and said labeled moiety.

20 In still yet another embodiment, the methods involve the steps of: (a) providing at least one nucleic acid coupled to a support; (b) providing a labeled moiety capable of being coupled with a ligase to said nucleic acid; (c) providing said ligase; and (d) coupling said labeled moiety to said nucleic acid using said ligase.

25 This invention also provides compounds of the formulas described herein.

Definitions

30 The following definitions are used, unless otherwise described. Halo is fluoro, chloro, bromo, or iodo. Alkyl, alkoxy, aralkyl, alkylaryl, etc. denote both straight and branched alkyl groups; but reference to an individual radical such as "propyl" embraces only the straight chain radical, a branched chain

isomer such as "isopropyl" being specifically referred to. Aryl includes a phenyl radical or an ortho-fused bicyclic carbocyclic radical having about nine to ten ring atoms in which at least one ring is aromatic. Heteroaryl encompasses a radical attached via a ring carbon of a monocyclic aromatic ring containing five or six ring atoms consisting of carbon and one to four heteroatoms each selected from the group consisting of non-peroxide oxygen, sulfur, and N(X) wherein X is absent or is H, O, (C₁-C₄)alkyl, phenyl or benzyl, as well as a radical of an ortho-fused bicyclic heterocycle of about eight to ten ring atoms derived therefrom, particularly a benz-derivative or one derived by fusing a propylene, trimethylene, or tetramethylene diradical thereto.

"Alkyl" refers to a straight chain, branched or cyclic chemical groups containing only carbon and hydrogen. Alkyl groups include, without limitation, ethyl, propyl, butyl, pentyl, cyclopentyl and 2-methylbutyl. Alkyl groups are unsubstituted or substituted with 1 or more substituents (e.g., halogen, alkoxy, amino).

"Alkylene" refers to a straight chain, branched or cyclic chemical group containing only carbon and hydrogen. Alkyl groups include, without limitation, ethylene, propylene, butylene, pentylene, and 2-methylbutylene. Alkyl groups are unsubstituted or substituted with 1 or more substituents (e.g., halogen, alkoxy, amino).

"Aryl" refers to a monovalent, unsaturated aromatic carbocyclic group. Aryl groups include, without limitation, phenyl, naphthyl, anthryl and biphenyl. Aryl groups are unsubstituted or substituted with 1 or more substituents (e.g. halogen, alkoxy, amino). "Arylene" refers to a divalent aryl group.

"Amido" refers to a chemical group having the structure $-C(O)NR_3-$, wherein R₃ is hydrogen, alkyl or aryl. Preferably, the amido group is of the structure $-C(O)NR_3-$ where R₃ is hydrogen or alkyl having from about 1 to about 6 carbon atoms. More preferably, the amido alkyl group is of the structure $-C(O)NH-$.

“Alkanoyl” refers to a chemical group having the structure $-(CH_2)_nC(O)-$, n is an integer ranging from 0 to about 10. Preferably, the alkanoyl group is of the structure $-(CH_2)_nC(O)-$, wherein n is an integer ranging from about 2 to about 10. More preferably, the alkanoyl group is of the structure $-(CH_2)_nC(O)-$, wherein n is an integer ranging from about 2 to about 6. Most preferably, the alkanoyl group is of the structure $-CH_2C(O)-$.

“Alkyl amido” refers to a chemical group having the structure $-R_4C(O)NR_3-$, wherein R_3 is hydrogen, alkyl or aryl, and R_4 is alkylene or arylene. Preferably, the alkyl amido group is of the structure $-(CH_2)_nC(O)NH-$, wherein n is an integer ranging from about 1 to about 10. More preferably, n is an integer ranging from about 1 to about 6. Most preferably, the alkyl amido group has the structure $-(CH_2)_2C(O)NH-$ or the structure $-CH_2C(O)NH-$.

“N-Amido alkyl” refers to a chemical group having the structure $-C(O)NR_3R_4-$, wherein R_3 is hydrogen, alkyl or aryl, and R_4 is alkylene or arylene. Preferably, the N-amido alkyl group is of the structure $-C(O)NH(CH_2)_nR_5-$, wherein n is an integer ranging from about 2 to about 10, and R_5 is O, NR_6 , or $C(O)$, and wherein R_6 is hydrogen, alkyl or aryl. More preferably, the N-amido alkyl group is of the structure $-C(O)NH(CH_2)_nN(H)-$, wherein n is an integer ranging from about 2 to about 6. Most preferably, the N-amido alkyl group is of the structure $-C(O)NH(CH_2)_4N(H)-$.

“Alkynyl alkyl” refers to a chemical group having the structure $-C\equiv C-R_4-$, wherein R_4 is alkyl or aryl. Preferably, the alkynyl alkyl group is of the structure $-C\equiv C-(CH_2)_nR_5-$, wherein n is an integer ranging from 1 to about 10, and R_5 is O, NR_6 or $C(O)$, wherein R_6 is hydrogen, alkyl or aryl. More preferably, the alkynyl alkyl group is of the structure $-C\equiv C-(CH_2)_nN(H)-$, wherein n is an integer ranging from 1 to about 4. Most preferably, the alkynyl alkyl group is of the structure $-C\equiv C-CH_2N(H)-$.

“Alkenyl alkyl” refers to a chemical group having the structure $-CH=CH-R_4-$, wherein R_4 is a bond, alkyl or aryl. Preferably, the alkenyl alkyl group is of the structure $-CH=CH-(CH_2)_nR_5-$, wherein n is an integer ranging from 0 to about 10, and R_5 is O, NR_6 , $C(O)$ or $C(O)NR_6$, wherein R_6 is

hydrogen, alkyl or aryl. More preferably, the alkenyl alkyl group is of the structure $-\text{CH}=\text{CH}-(\text{CH}_2)_n\text{C}(\text{O})\text{NR}_6-$, wherein n is an integer ranging from 0 to about 4. Most preferably, the alkenyl alkyl group is of the structure $-\text{CH}=\text{CH}-\text{C}(\text{O})\text{N}(\text{H})-$.

5 “Functionalized alkyl” refers to a chemical group of the structure $-(\text{CH}_2)_n\text{R}_7-$, wherein n is an integer ranging from 1 to about 10, and R_7 is O, S, NH or C(O). Preferably, the functionalized alkyl group is of the structure $-(\text{CH}_2)_n\text{C}(\text{O})-$, wherein n is an integer ranging from 1 to about 4. More preferably, the functionalized alkyl group is of the structure $-\text{CH}_2\text{C}(\text{O})-$.

10 “Alkoxy” refers to a chemical group of the structure $-\text{O}(\text{CH}_2)_n\text{R}_8-$, wherein n is an integer ranging from 2 to about 10, and R_8 is a bond, O, S, NH or C(O). Preferably, the alkoxy group is of the structure $-\text{O}(\text{CH}_2)_n-$, wherein n is an integer ranging from 2 to about 4. More preferably, the alkoxy group is of the structure $-\text{OCH}_2\text{CH}_2-$.

15 “Alkyl thio” refers to a chemical group of the structure $-\text{S}(\text{CH}_2)_n\text{R}_8-$, wherein n is an integer ranging from 1 to about 10, and R_8 is a bond, O, S, NH or C(O). Preferably, the alkyl thio group is of the structure $-\text{S}(\text{CH}_2)_n-$, wherein n is an integer ranging from 2 to about 4. More preferably, the thio group is of the structure $-\text{SCH}_2\text{CH}_2\text{C}(\text{O})-$.

20 “Amino alkyl” refers to a chemical group having an amino group attached to an alkyl group. Preferably an amino alkyl is of the structure $-(\text{CH}_2)_n\text{NH}-$, wherein n is an integer ranging from about 2 to about 10. More preferably it is of the structure $-(\text{CH}_2)_n\text{NH}-$, wherein n is an integer ranging from about 2 to about 4. Most preferably, the amino alkyl group is of the
25 structure $-(\text{CH}_2)_2\text{NH}-$.

 “Nucleic acid” refers to a polymer comprising 2 or more nucleotides and includes single-, double- and triple stranded polymers. “Nucleotide” refers to both naturally occurring and non-naturally occurring compounds and comprises a heterocyclic base, a sugar, and a linking group, preferably a
30 phosphate ester. For example, structural groups may be added to the ribosyl or deoxyribosyl unit of the nucleotide, such as a methyl or allyl group at the 2'-O

position or a fluoro group that substitutes for the 2'-O group. The linking group, such as a phosphodiester, of the nucleic acid may be substituted or modified, for example with methyl phosphonates or O-methyl phosphates. Bases and sugars can also be modified, as is known in the art. "Nucleic acid," for the purposes of
5 this disclosure, also includes "peptide nucleic acids" in which native or modified nucleic acid bases are attached to a polyamide backbone.

The phrase "coupled to a support" means bound directly or indirectly thereto including attachment by covalent binding, hydrogen bonding, ionic interaction, hydrophobic interaction, or otherwise.

10 "Probe" refers to a nucleic acid that can be used to detect, by hybridization, a target nucleic acid. Preferably, the probe is complementary to the target nucleic acid along the entire length of the probe, but hybridization can occur in the presence of one or more base mismatches between probe and target.

"Perfect match probe" refers to a probe that has a sequence that is
15 perfectly complementary to a particular target sequence. The test probe is typically perfectly complementary to a portion (subsequence) of the target sequence. The perfect match (PM) probe can be a "test probe", a "normalization control" probe, an expression level control probe and the like. A perfect match control or perfect match probe is, however, distinguished from a "mismatch
20 control" or "mismatch probe." In the case of expression monitoring arrays, perfect match probes are typically preselected (designed) to be complementary to particular sequences or subsequences of target nucleic acids (e.g., particular genes). In contrast, in generic difference screening arrays, the particular target sequences are typically unknown. In the latter case, perfect match probes cannot
25 be preselected. The term perfect match probe in this context is to distinguish that probe from a corresponding "mismatch control" that differs from the perfect match in one or more particular preselected nucleotides as described below.

"Mismatch control" or "mismatch probe", in expression monitoring arrays, refers to probes whose sequence is deliberately selected not to be
30 perfectly complementary to a particular target sequence. For each mismatch (MM) control in a high-density array there preferably exists a corresponding perfect match (PM) probe that is perfectly complementary to the same particular

target sequence. In "generic" (*e.g.*, random, arbitrary, haphazard, *etc.*) arrays, since the target nucleic acid(s) are unknown perfect match and mismatch probes cannot be *a priori* determined, designed, or selected. In this instance, the probes are preferably provided as pairs where each pair of probes differ in one or more preselected nucleotides. Thus, while it is not known *a priori* which of the probes in the pair is the perfect match, it is known that when one probe specifically hybridizes to a particular target sequence, the other probe of the pair will act as a mismatch control for that target sequence. It will be appreciated that the perfect match and mismatch probes need not be provided as pairs, but may be provided as larger collections (*e.g.*, 3, 4, 5, or more) of probes that differ from each other in particular preselected nucleotides. While the mismatch(s) may be located anywhere in the mismatch probe, terminal mismatches are less desirable as a terminal mismatch is less likely to prevent hybridization of the target sequence. In a particularly preferred embodiment, the mismatch is located at or near the center of the probe such that the mismatch is most likely to destabilize the duplex with the target sequence under the test hybridization conditions. In a particularly preferred embodiment, perfect matches differ from mismatch controls in a single centrally-located nucleotide.

"Labeled moiety" refers to a moiety capable of being detected by the various methods discussed herein or known in the art.

Nucleic Acid Labeling Compounds

The nucleic acid labeling compounds of the present invention are of the following structure:



wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid; T is an optional template moiety; H_c is an heterocyclic group; L is a linker moiety; Q is a detectable moiety; and M is an connecting group, wherein m is an integer ranging from 0 to about 20. In a preferred embodiment m is from 0 to about 15. In a more preferred embodiment m is from 0 to about 10. In the most preferred embodiment m is from 0 to about 5.

The group A is either hydrogen or a functional group that permits the attachment of a nucleic acid labeling compound to a nucleic acid. Nonlimiting examples of such groups include the following: monophosphate; diphosphate; triphosphate (H_4O_9P); phosphoramidite $((R_2N)(R'O)P)$, wherein R is linear, 5 branched or cyclic alkyl, and R' is a protecting group such as 2-cyanoethyl; and H-phosphonate $(HP(O)O-HNR_3)$, wherein R is linear, branched or cyclic alkyl.

The template moiety (T) is covalently attached to a methylene group (CH_2) at one position and a heterocyclic group (H_c) or the linker moiety at another position. A nonlimiting set of template moieties is shown in FIG. 1, 10 wherein the substituents are defined as follows: X is O, S, NR_1 or CHR_2 ; Y is H, N_3 , F, OR_9 , SR_9 or NHR_9 ; Z is H, N_3 , F or OR_{10} ; W is O, S or CH_2 ; D is O or S; and, G is O, NH or CH_2 . The substituents R_1 , R_2 , R_9 and R_{10} are independent of one another and are H, alkyl or aryl.

The heterocyclic group (H_c) is a cyclic moiety containing both carbon 15 and a heteroatom.

The linker moiety (L) of the nucleic acid labeling compound is covalently bound to the heterocycle (H_c) or the template moiety at one terminal position. It is attached to the detectable moiety (Q) at another terminal position, either directly or through a connecting group (M). It is of a structure that is 20 sterically and electronically suitable for incorporation into a nucleic acid. Nonlimiting examples of linker moieties include amido alkyl groups, functionalized alkyl groups, alkenyl alkyl groups, alkanoyl groups, and N-alkyl amido groups.

Amido groups have the structure $-C(O)NR_3-$, wherein R_3 is 25 hydrogen, alkyl or aryl. Preferably, the amido group is of the structure $-C(O)NR_3-$ where R_3 is hydrogen or alkyl having from about 1 to about 6 carbon atoms. More preferably, the amido alkyl group is of the structure $-C(O)NH-$.

Alkanoyl groups have the structure $-(CH_2)_nC(O)-$, n is an integer 30 ranging from 0 to about 10. Preferably, the alkanoyl group is of the structure $-(CH_2)_nC(O)-$, wherein n is an integer ranging from about 2 to about 10. More

preferably, the alkanoyl group is of the structure $-(CH_2)_nC(O)-$, wherein n is an integer ranging from about 2 to about 6. Most preferably, the alkanoyl group is of the structure $-CH_2C(O)-$.

Amido alkyl groups have the structure $-R_4C(O)NR_3-$, wherein R_3 is hydrogen, alkyl or aryl, and R_4 is alkyl or aryl. The amido alkyl group is preferably of the structure $-(CH_2)_nC(O)NH-$, wherein n is an integer ranging from about 1 to about 10. More preferably, n is an integer ranging from about 1 to about 6. Most preferably, the alkyl amido group has the structure $-(CH_2)_2C(O)NH-$ or the structure $-CH_2C(O)NH-$.

N-Amido alkyl groups have the structure $-C(O)NR_3R_4-$, wherein R_3 is hydrogen, alkyl or aryl, and R_4 is alkylene or arylene. Preferably, the N-amido alkyl group is of the structure $-C(O)NH(CH_2)_nR_5-$, wherein n is an integer ranging from about 2 to about 10, and R_5 is O, NR_6 , or $C(O)$, and wherein R_6 is hydrogen, alkyl or aryl. More preferably, the N-amido alkyl group is of the structure $-C(O)NH(CH_2)_nN(H)-$, wherein n is an integer ranging from about 2 to about 6. Most preferably, the N-amido alkyl group is of the structure $-C(O)NH(CH_2)_4N(H)-$.

Alkenyl alkyl groups are of the structure $-CH=CH-R_4-$, wherein R_4 is alkyl or aryl. The alkenyl alkyl group is preferably of the structure $-CH=CH(CH_2)_nR_5-$, wherein n is an integer ranging from 1 to about 10, and R_5 is O, NR_6 , $C(O)$ or $C(O)NR_6$, wherein R_6 is hydrogen, alkyl or aryl. More preferably, the alkenyl alkyl group is of the structure $-CH=CH-(CH_2)_nC(O)NR_6-$, wherein n is an integer ranging from 0 to about 4. Most preferably, the alkenyl alkyl group is of the structure $-CH=CH-C(O)N(H)-$.

Functionalized alkyl groups are of the structure $-(CH_2)_nR_7-$, wherein n is an integer ranging from 1 to about 10, and R_7 is O, S, NH, or $C(O)$. The functionalized alkyl group is preferably of the structure $-(CH_2)_nC(O)-$, wherein n is an integer ranging from 1 to about 4. More preferably, the functionalized alkyl group is $-CH_2C(O)-$.

Alkoxy groups are of the structure $-(CH_2)_nR_8-$, wherein n is an integer ranging from 2 to about 10, and R_8 is O, S, NH, or C(O). The alkoxy group is preferably of the structure $-(CH_2)_n-$ or $-(CH_2)_nC(O)-$, wherein n is an integer ranging from 2 to about 4. More preferably, the alkoxy group is of the structure $-OCH_2CH_2-$ or $-OCH_2CH_2C(O)-$.

Alkyl thio groups are of the structure $-(CH_2)_nR_8-$, wherein n is an integer ranging from 2 to about 10, and R_8 is O, S, NH, or C(O). The alkyl thio group is preferably of the structure $-(CH_2)_nC(O)-$, wherein n is an integer ranging from 2 to about 4. More preferably, the alkyl thio group is of the structure $-SCH_2CH_2C(O)-$.

Amino alkyl groups comprise an amino group attached to an alkyl group. Preferably, amino alkyl groups are of the structure $-NH(CH_2)_nNH-$, wherein n is an integer ranging from about 2 to about 10. The amino alkyl group is more preferably of the structure $-(CH_2)_nNH-$, wherein n is an integer ranging from about 2 to about 4. Most preferably, the amino alkyl group is of the structure $-(CH_2)_4NH-$.

The detectable moiety (Q) is a chemical group that provides an signal. The signal is detectable by any suitable means, including spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. In certain cases, the signal is detectable by 2 or more means.

The detectable moiety provides the signal either directly or indirectly. A direct signal is produced where the labeling group spontaneously emits a signal, or generates a signal upon the introduction of a suitable stimulus. Radiolabels, such as 3H , ^{125}I , ^{35}S , ^{14}C or ^{32}P , and magnetic particles, such as Dynabeads™, are nonlimiting examples of groups that directly and spontaneously provide a signal. Labeling groups that directly provide a signal in the presence of a stimulus include the following nonlimiting examples: colloidal gold (40 - 80 nm diameter), which scatters green light with high efficiency; fluorescent labels, such as fluorescein, texas red, rhodamine, and green fluorescent protein (Molecular Probes, Eugene, Oregon), which absorb and subsequently emit light; chemiluminescent or bioluminescent labels, such as

luminol, lophine, acridine salts and luciferins, which are electronically excited as the result of a chemical or biological reaction and subsequently emit light; spin labels, such as vanadium, copper, iron, manganese and nitroxide free radicals, which are detected by electron spin resonance (ESR) spectroscopy; dyes, such as
5 quinoline dyes, triarylmethane dyes and acridine dyes, which absorb specific wavelengths of light; and colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads. See U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241.

A detectable moiety provides an indirect signal where it interacts
10 with a second compound that spontaneously emits a signal, or generates a signal upon the introduction of a suitable stimulus. Biotin, for example, produces a signal by forming a conjugate with streptavidin, which is then detected. See Hybridization With Nucleic Acid Probes. In *Laboratory Techniques in Biochemistry and Molecular Biology*; Tijssen, P., Ed.; Elsevier: New York,
15 1993; Vol. 24. An enzyme, such as horseradish peroxidase or alkaline phosphatase, that is attached to an antibody in a label-antibody-antibody as in an ELISA assay, also produces an indirect signal.

A preferred detectable moiety is a fluorescent group. Fluorescent groups typically produce a high signal to noise ratio, thereby providing increased
20 resolution and sensitivity in a detection procedure. Preferably, the fluorescent group absorbs light with a wavelength above about 300 nm, more preferably above about 350 nm, and most preferably above about 400 nm. The wavelength of the light emitted by the fluorescent group is preferably above about 310 nm, more preferably above about 360 nm, and most preferably above about 410 nm.

25 The fluorescent detectable moiety is selected from a variety of structural classes, including the following nonlimiting examples: 1- and 2-aminonaphthalene, p,p'-diaminostilbenes, pyrenes, quaternary phenanthridine salts, 9-aminoacridines, p,p'-diaminobenzophenone imines, anthracenes, oxacarbocyanine, marocyanine, 3-aminoequilenin, perylene, bisbenzoxazole,
30 bis-p-oxazolyl benzene, 1,2-benzophenazin, retinol, bis-3-aminopridinium salts, hellebrigenin, tetracycline, sterophenol, benzimidazolyl phenylamine, 2-oxo-3-chromen, indole, xanthen, 7-hydroxycoumarin, phenoxazine, salicylate, strophanthidin, porphyrins, triarylmethanes, flavin, xanthene dyes (e.g.,

fluorescein and rhodamine dyes); cyanine dyes; 4,4-difluoro-4-bora-3a,4a-diazas-indacene dyes and fluorescent proteins (e.g., green fluorescent protein, phycobiliprotein).

A number of fluorescent compounds are suitable for incorporation into the present invention. Nonlimiting examples of such compounds include the following: dansyl chloride; fluoresceins, such as 3,6-dihydroxy-9-phenylxanthhydrol; rhodamineisothiocyanate; N-phenyl-1-amino-8-sulfonatonaphthalene; N-phenyl-2-amino-6-sulfonatonaphthalene; 4-acetamido-4-isothiocyanatostilbene-2,2'-disulfonic acid; pyrene-3-sulfonic acid; 2-toluidinonaphthalene-6-sulfonate; N-phenyl, N-methyl 2-aminonaphthalene-6-sulfonate; ethidium bromide; stebrine; auromine-0,2-(9'-anthroyl)palmitate; dansyl phosphatidylethanolamin; N,N'-dioctadecyl oxacarbocyanine; N,N'-dihexyl oxacarbocyanine; merocyanine, 4-(3'-pyrenyl)butyrate; d-3-aminodesoxy-equilenin; 12-(9'-anthroyl)stearate; 2-methylanthracene; 9-vinylanthracene; 2,2'-(vinylene-p-phenylene)bisbenzoxazole; p-bis[2-(4-methyl-5-phenyl oxazolyl)]benzene; 6-dimethylamino-1,2-benzophenazin; retinol; bis(3'-aminopyridinium)-1,10-decandiyl diiodide; sulfonaphthylhydrazone of hellbrienin; chlorotetracycline; N-(7-dimethylamino-4-methyl-2-oxo-3-chromenyl)maleimide; N-[p-(2-benzimidazolyl)phenyl]maleimide; N-(4-fluoranthyl)maleimide; bis(homovanillic acid); resazarin; 4-chloro-7-nitro-2,1,3-benzooxadizole; merocyanine 540; resorufin; rose bengal and 2,4-diphenyl-3(2H)-furanone. Preferably, the fluorescent detectable moiety is a fluorescein or rhodamine dye.

Another preferred detectable moiety is colloidal gold. The colloidal gold particle is typically 40 to 80 nm in diameter. The colloidal gold may be attached to a labeling compound in a variety of ways. In one embodiment, the linker moiety of the nucleic acid labeling compound terminates in a thiol group (-SH), and the thiol group is directly bound to colloidal gold through a dative bond. See Mirkin et al. *Nature* **1996**, 382, 607-609. In another embodiment, it is attached indirectly, for instance through the interaction between colloidal gold conjugates of antibiotin and a biotinylated labeling compound. The detection of the gold labeled compound may be enhanced through the use of a silver enhancement method. See Danscher et al. *J. Histotech* **1993**, 16, 201-207.

The connecting groups (M)_m may serve to covalently attach the linker group (L) to the detectable moiety (Q). Each M group can be the same or different and can independently be any suitable structure that will not interfere with the function of the labeling compound. Nonlimiting examples of M groups include the following: amino alkyl, -CO(CH₂)₅NH-, -CO-, -NH-, -CO(O)-, -CO(NH)-, -(CH₂)_iO-, -(CH₂)_jNH-, -C(O)(CH₂)_hO-, -CO(CH₂)₅NHCO(CH₂)₅NH-, -NH(CH₂CH₂O)_kNH-, and -CO(CH₂)₅-; wherein, k is an integer from 1 to about 5, preferably k is 1 or 2; m is an integer ranging from 0 to about 5, preferably 0 to about 3; h, i and j are independently integers from 1 to about 5, preferably 1 to about 3.

Synthesis of Nucleic Acid Labeling Compounds

Figure 2 illustrates a synthetic route to nucleic acid labeling compound, 4, starting from 2-amino-5-(β-D-ribofuranosyl)-4(1H)-pyrimidinone, 1. The hydroxy groups are protected as the trimethyl silylethers according to the procedure described in *J. Antibiot.* 1977, 30, 129. The N1 nitrogen atom is alkylated using a substituted iodoamide having a biotin label attached, according to the procedure described in *Tetrahedron* 1984, 40, 33. (PEO-iodoacetyl biotin can be purchased from Pierce Chemical Co.) The silyl protecting groups are removed, providing a compound having free alcohol groups using acetic acid in methanol. The product is converted into a 5'-triphosphate to afford labeled compound 4.

Figures 3 illustrates a synthetic route to nucleic acid labeling compounds, 6 and 7, starting from 2-amino-5-(β-D-ribofuranosyl)-4(1H)-pyrimidinone, 1. The hydroxy groups are protected as the trimethyl silylethers according to the procedure described in *J. Antibiot.* 1977, 30, 129. The N1 nitrogen atom is alkylated according to the procedure described in *Tetrahedron* 1984, 40, 33, using ethyl iodoacetate. The ester is converted to amide, 5, using ethylene diamine. The biotin labeled compound is prepared by reacting amide, 5, with biotin-dPEG4TM-NHS (purchased from Quanta Biodesign). Alternatively, amide 5 is reacted with 5-carboxyfluorescein-NHS. The silyl protecting groups are removed to provide a compound having free alcohol groups, using acetic acid in methanol. The product is converted into a 5'-

triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compounds 6 and 7.

Figure 4 illustrates an alternate synthetic route to nucleic acid labeling compounds, 6 and 7, starting from 2-amino-5-(β -D-ribofuranosyl)-4(1*H*)-pyrimidinone, 1. The hydroxy groups are protected as acetyl esters using acetic anhydride in pyridine. The N1 nitrogen atom is alkylated according to the procedure described in *Tetrahedron* 1984, 40, 33, using ethyl iodoacetate. The ester is converted to amide, 8, using ethylene diamine. The biotin labeled compound is prepared by reacting amide, 8, with biotin-dPEG4TM-NHS (purchased from Quanta Biodesign). Alternatively, amide 8 is reacted with 5-carboxyfluorescein-NHS. The silyl protecting groups are removed to provide a compound having free alcohol groups, using acetic acid in methanol. The product is converted into a 5'-triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compounds 6 and 7.

Figure 5 illustrates a synthetic route to nucleic acid labeling compounds, 10 and 11, starting from 2-amino-5-(β -D-ribofuranosyl)-4(1*H*)-pyrimidinone, 1. The hydroxy groups are protected as acetyl esters using acetic anhydride in pyridine. The N1 nitrogen atom is alkylated according to the procedure described in K. Muhlegger, *et al.* 1996, WO 96/28640, using methyl acrylate in triethyl amine. The ester, 8 is converted to amide, 9, using ethylene diamine. The biotin labeled compound is prepared by reacting amide, 9, with biotin-dPEG4TM-NHS (purchased from Quanta Biodesign). Alternatively, amide 9 is reacted with 5-carboxyfluorescein-NHS. The product is converted into a 5'-triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compounds 10 and 11.

Figure 6 illustrates a synthetic route to nucleic acid labeling compounds, 14 and 15, starting from 2-amino-5-(β -D-ribofuranosyl)-4(1*H*)-pyrimidinone, 1. The hydroxy groups are protected as acetyl esters using acetic anhydride in pyridine. The N1 nitrogen atom is alkylated according to the procedure described in *Tet Lett* 1995 36, 3261, using methyl propynoate. The ester, 12 is converted to amide, 9, using ethylene diamine. The biotin labeled compound is prepared by reacting amide, 9, with biotin-dPEG4TM-NHS

(purchased from Quanta Biodesign). Alternatively, amide **9** is reacted with 5-carboxyfluorescein-NHS. The product is converted into a 5'-triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compounds **14** and **15**.

5 Figure 7 illustrates a synthetic route to nucleic acid labeling compounds, **17** and **18**, starting from 2-amino-5-(β -D-ribofuranosyl)-4(1*H*)-pyrimidinone, **1**. The hydroxy groups are protected as the trimethyl silylethers according to the procedure described in *J. Antibiot.* **1977**, *30*, 129. The N1 nitrogen atom is alkylated, using 2-(2-chloromethoxy-ethyl)-isoindole-1,3-dione, 10 according to the procedure described in *Chemo* **1985**, *31*, 151. The product, **16** is treated with hydrazine in ethanol to remove the protecting groups. The biotin labeled compound is prepared by reacting the ether-amine with biotin-dPEG4TM-NHS (purchased from Quanta Biodesign) using standard conditions. Alternatively, the ether-amine is reacted with (Ac)₂Fl-NHS using standard 15 conditions. The product is converted into a 5'-triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compounds **17** and **18**.

 Figure 8 illustrates a synthetic route to nucleic acid labeling compounds, **20** and **21**, starting from 2-amino-5-(β -D-ribofuranosyl)-4(1*H*)- 20 pyrimidinone, **1**. The hydroxy groups are protected as acetyl esters using acetic anhydride in pyridine. The N1 nitrogen atom is alkylated using (6-trifluoroacetamido)hexyl alcohol in the presence of triphenyl phosphane ((Ph)₃P), and diethylazodicarboxylate (DEAD), according to the procedure described in Brossette, T. *et al.*, *J. Org. Chem.* **1999**, *64*, 5083 and *Nucleosides, Nucleotides and* 25 *Nucleic Acids* **2000**, *19*, 867. The N1 nitrogen atom is alkylated using (6-trifluoroacetamido)hexyl alcohol according to the procedure described in *J. Antibiot.* **1977**, *30*, 129. The protecting groups are removed to provide a compound having free amine and alcohol groups, using ammonia in methanol. The biotin labeled compound is prepared by reacting the ether-amine with 30 biotin-dPEG4TM-NHS (purchased from Quanta Biodesign) using standard conditions. Alternatively, the ether-amine is reacted with (Ac)₂Fl-NHS using standard conditions. The product is converted into a 5'-triphosphate using

standard phosphorylation conditions to afford, respectively, nucleic acid labeling compounds **20** and **21**.

Figure 9 illustrates a synthetic route to nucleic acid labeling compounds, **23** and **24**, starting from 2-amino-5-(β -D-ribofuranosyl)-4(1*H*)-pyrimidinone, **1**. The hydroxy groups are protected as acetyl esters using acetic anhydride in pyridine. The N1 nitrogen atom is alkylated using 2-(6-iodohexyl)-isoindole-1,3-dione according to the procedure described in *J. Antibiot.* **1977**, *30*, 129. The product, **22** is treated with hydrazine in ethanol to remove the protecting groups. The biotin labeled compound is prepared by reacting the free amine with biotin-dPEG4TM-NHS (purchased from Quanta Biodesign) using standard conditions. Alternatively, the ether-amine is reacted with (Ac)₂Fl-NHS using standard conditions. The product is converted into a 5'-triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compounds **23** and **24**.

Figure 10 illustrates a synthetic route to nucleic acid labeling compounds, **36**, starting from [1,4]Dioxan-2-one, Shahi, S.P., *et al.*, *J. Org. Chem.* (1999), *64*: 4509-11, Nishimura, T., *et al.*, *J. Org. Chem.* **1999**, *64*: 6750-55, Nishida, H., *et al.*, *J. Polym. Sci.* **2000**, *38*: 1560-67; (2-Hydroxy-ethoxy)-acetic acid methyl ester, Kitano, M.; and Ohashi, N., EP 787728 A1(1997), or a polymer thereof, Nishida H, *et al. J. Polym. Sci.*, (2000), *38*: 1560-67, **33**. The compound is reacted with a diamine having a suitable linker moiety, *e.g.*, a polyethylene oxide, alkylene, a combination thereof and the like. The labeled compound is prepared by reacting the ether-amine with a suitable labeling compound using standard conditions. The product is converted into a 5'-triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compound **36**.

Nucleic Acid Labeling

Nucleic acids can be isolated from a biological sample or synthesized, on a solid support or in solution for example, according to methods known to those of skill in the art. As used herein, there is no limitation on the length or source of the nucleic acid used in a labeling process. Exemplary methods of nucleic acid isolation and purification are described in Theory and

Nucleic Acid Preparation. In *Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes*; P. Tijssen, Ed.; Part I; Elsevier: N.Y., 1993. A preferred method of isolation involves an acid guanidinium-phenol-chloroform extraction followed by oligo dT column
5 chromatography or (dT)_n magnetic bead use. Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Laboratory, 1989; Vols. 1-3; and *Current Protocols in Molecular Biology*; F. Ausubel et al. Eds.; Greene Publishing and Wiley Interscience: N.Y., 1987.

In certain cases, the nucleic acids are increased in quantity through
10 amplification. Suitable amplification methods include, but are not limited to, the following examples: polymerase chain reaction (PCR) (Innis, et al. *PCR Protocols. A guide to Methods and Application*; Academic Press: San Diego, 1990); ligase chain reaction (LCR) (Wu and Wallace. *Genomics* 1989, 4, 560; Landgren, et al. *Science* 1988, 241, 1077; and Barringer, et al. *Gene* 1990, 89,
15 117); transcription amplification (Kwoh et al. *Proc. Natl. Acad. Sci. USA* 1989, 86, 1173); and self-sustained sequence replication (Guatelli, et al. *Proc. Nat. Acad. Sci. USA* 1990, 87, 1874).

The nucleic acid labeling compound can be incorporated into a nucleic acid using a number of methods. For example, it can be directly attached
20 to an original nucleic acid sample (e.g., mRNA, polyA mRNA, cDNA) or to an amplification product. Methods of attaching a labeling compound to a nucleic acid include, without limitation, nick translation, 3-end-labeling, ligation, in vitro transcription (IVT) or random priming. Where the nucleic acid is an RNA, a labeled riboligonucleotide is ligated, for example, using an RNA ligase such
25 as T4 RNA Ligase. In *The Enzymes*; Uhlenbeck and Greensport, Eds.; Vol. XV, Part B, pp. 31-58; and, Sambrook et al., pp. 5.66-5.69. Terminal transferase is used to add deoxy-, dideoxy- or ribonucleoside triphosphates (dNTPs, ddNTPs or NTPs), for example, where the nucleic acid is single stranded DNA.

The labeling compound can also be incorporated at an internal
30 position of a nucleic acid. For example, PCR in the presence of a labeling compound provides an internally labeled amplification product. See, e.g., Yu et al. *Nucleic Acids Research* 1994, 22, 3226-3232. Similarly, IVT in the presence of a labeling compound can provide an internally labeled nucleic acid.

Probe Hybridization

The nucleic acid to which the labeling compound is attached can be detected after hybridization with a nucleic acid probe. Alternatively, the probe can be labeled, depending upon the experimental scheme preferred by the user.

- 5 The probe is a nucleic acid, or a modified nucleic acid, that is either attached to a solid support or is in solution. It is complementary in structure to the labeled nucleic acid with which it hybridizes. The solid support is of any suitable material, including polystyrene based beads and glass chips. In a preferred embodiment, the probe or target nucleic acid is attached to a glass chip, such as a
- 10 GeneChip[®] product (Affymetrix, Inc., Santa Clara, CA). See International Publication Nos. WO 97/10365, WO 97/29212, WO 97/27317, WO 95/11995, WO 90/15070, and U.S. Pat. Nos. 5,744,305 and 5,445,934 which are hereby incorporated by reference.

- Because probe hybridization is often a step in the detection of a
- 15 nucleic acid, the nucleic acid labeling compound must be of a structure that does not substantially interfere with that process. The steric and electronic nature of the labeling compound, therefore, is compatible with the binding of the attached nucleic acid to a complementary structure.

20

EXAMPLES

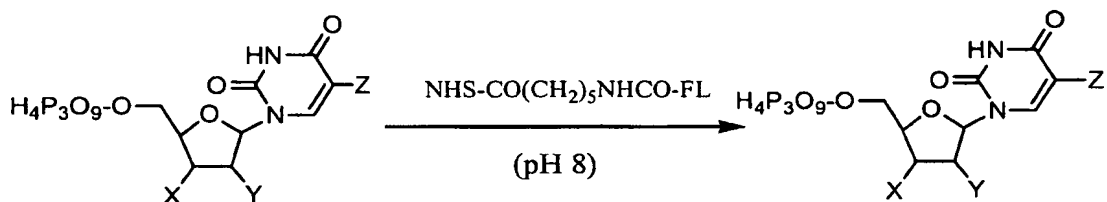
The following examples are offered to illustrate, but not to limit, the present invention.

General Experimental Details

- Reagents were purchased from Aldrich Chemical Company
- 25 (Milwaukee, WI) in the highest available purity. All listed solvents were anhydrous. Intermediates were characterized by ¹H NMR and mass spectrometry.

Example 1Synthesis of fluorescein derivatives of 2'-amino-2'-deoxyuridine triphosphate and 3'-amino-3'-deoxythymidinetriphosphate (Scheme 1)

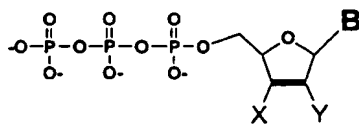
To 0.5 umoles (50 uL of a 10 mM solution) of the amino nucleotide triphosphate (1 or 2) in a 0.5 mL eppendorf tube was added 25 uL of a 1 M aqueous solution of sodium borate, pH 8, 87 uL of methanol, and 88 uL (10 umol, 20 equiv) of a 100 mM solution of 5-carboxyfluorescein -X-NHS ester in methanol. The mixture was vortexed briefly and allowed to stand at room temperature in the dark for 15 hours. The sample was then purified by ion-exchange HPLC to afford the fluoresceinated derivatives 3 or 4 in about 78-84% yield. Relative efficiencies of incorporation of these compounds by TdT are shown in Table 1.

Scheme 1

99 X = OH, Y = NH₂, Z = H
 98 X = NH₂, Y = H, Z = CH₃

97 X = OH, Y = NHCO(CH₂)₅NHCOFL, Z = H
 96 X = NHCO(CH₂)₅NHCOFL, Y = H, Z = CH₃

15

Table 1.Incorporation of triphosphate compounds by TdT.

TdT Labeling Efficiencies			% Labeled	
X (3')	Y (2')	B (1'b)	40U	160U
OH	H	uracil	100.0	100.0

NH ₂	H	thymine	100	100
NHCO(CH ₂) ₅ NH-(CO-FL)	H	thymine	1.3	2.2
OH	NH ₂	uracil	65	95
OH	NHCO(CH ₂) ₅ NH-(CO-FL)	uracil	3.0	6.6
OH	O(CH ₂) ₆ NH-(CO-FL)	uracil	2.5	7.0
OH	O(CH ₂) ₆ NHCO-(CH ₂) ₅ -NHCO-Biotin	uracil	15.0	17.0
OH	NH(CH ₂) ₅ CH ₃	uracil	4.5	5.0

Example 2

5 Synthesis of N1-labeled 2-Amino-5-(β -D-ribofuranosyl)-4(1*H*)-pyrimidinone

Synthesis of the N1-labeled 2-amino-5-(β -D-ribofuranosyl)-4(1*H*)-pyrimidinone, **55**, involved alkylation at N1 using conditions similar to those described by Muehlegger, et al. (WO 96/28640) for the N1-alkylation of pyrazalo-[4,3-*d*]pyrimidines (Figure 3A).

10

Example 3

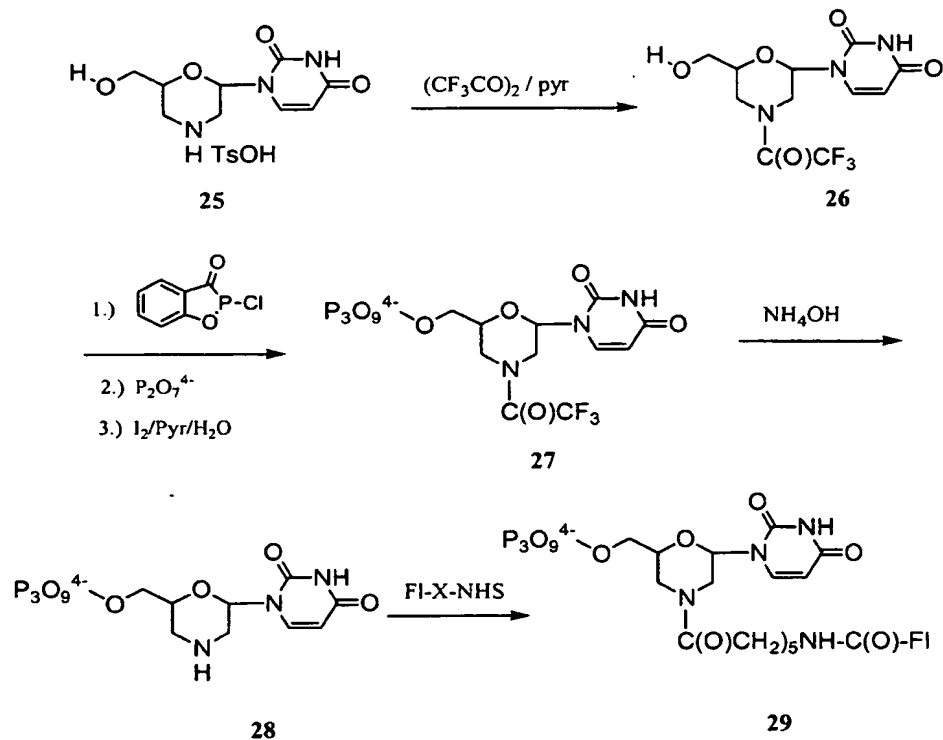
Synthesis of N-(6-(fluorescein-5-carboxamido)hexanoyl)-morpholino uridine triphosphate (Scheme 2).

Morpholino-uracil tosylate salt **25** (30 mg) was co-evaporated with pyridine (3x3 ml) and dissolved in 2 ml of pyridine and cooled to 0°C.

- 15 Trifluoroacetic anhydride (30 μ L) was added and stirred for 1 hour. The reaction was followed by HPLC until complete. The pyridine was removed and the residue was dissolved in 1 ml of water and filtered. The product was purified by HPLC on a Waters C-18 bondapak cartridge (Buffer : A= 50 mM TEM pH 7.0; B= acetonitrile) using a gradient of 0-25%B in 30 minutes (retention time = 22
- 20 min.). The product was desalted on a Sep-Pak cartridge and freeze-dried to give 151 mg of **26**. Phosphorylation of **2** using the POCl₃ method gave **27**. The removal of the trifluoroacetyl group with conc. NH₄OH at 50°C for 30 min to provide the free amine, **28**. Conjugation of **28** to 5-carboxyfluorescein-aminocaproic acid N-hydroxysuccinimide (Fl-X-NHS) under standard

conditions provided amide **29**. The mass spectral and NMR data for compounds **25-29** were consistent with the proposed structures.

Scheme 2



5

Example 4

The compounds of the invention having general formulas I, IA IB, II, III, IIIA, IIIB, and IV as disclosed herein can be prepared using available reagents and procedures that are standard in the art. In addition, examples of these procedures are illustrated in Figures 2-10.

Example 5

Procedure for HPLC Analysis of Enzymatic Incorporation of Modified Nucleotides.

Reaction Conditions

15 TdT

3 uM dT16 template

15(30) uM NTP
40 U TdT (Promega)
1X buffer, pH 7.5 (Promega)

Procedure: incubate 1hr. at 37 °C, then for 10 min. at 70 °C ,
5 followed by the addition of EDTA (2 mM final concentration) in a volume of 50 uL

HPLC Analysis

Materials and Reagents

10 4.6 mm X 250 mm Nucleopac PA-100 ion-exchange column
(Dionex) buffer A: 20 mM NaOH (or 20 mM Tris pH 8, in the case of TdT
incorporation of nucleotide triphosphates that are not dye-labeled) buffer B:
20mM NaOH, 1M NaCl (or 20 mM Tris pH 8, 1M NaCl, in the case of TdT
incorporation of nucleotide triphosphates that are not dye-labeled)

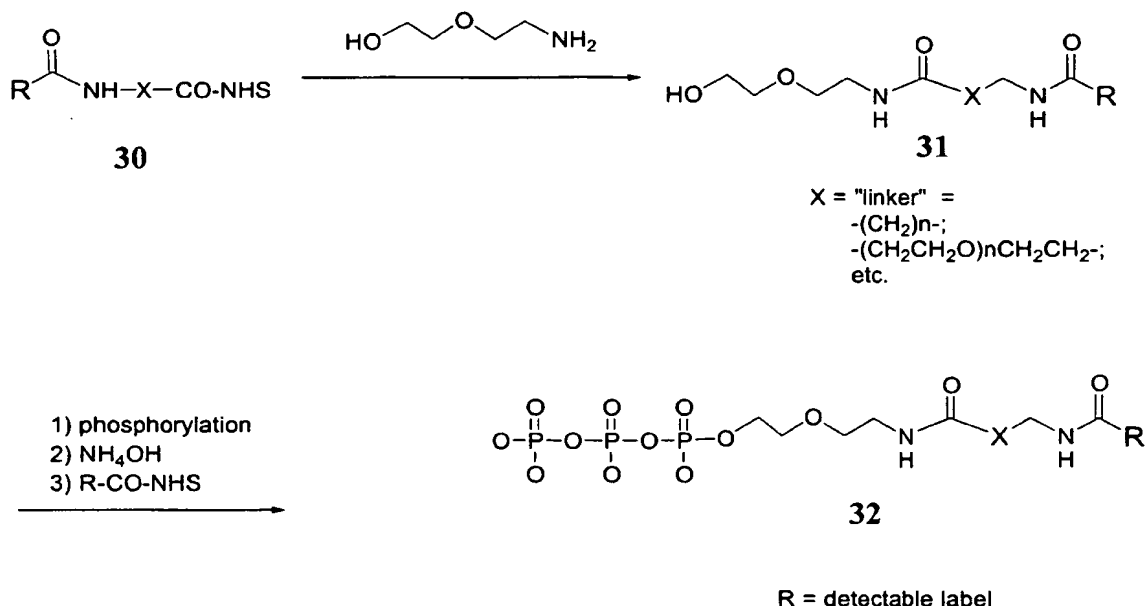
15 General Procedure

Dilute the reaction with 50 uL of buffer A. Inject 50 uL of this
sample onto the HPLC column and fractionate using a gradient of 5 to 100%
buffer B over 30 minutes at a flow rate of 1 mL/min.. Detect the peaks
simultaneously at 260 nm absorbance and the absorbance maximum of the dye
20 (or the fluorescence emission maximum of the dye).

The incorporation efficiency is expressed as the fraction of
oligonucleotide that is labeled. This number is determined by dividing the peak
area measured at 260 nm absorbance of the labeled oligonucleotide by the sum
of the peak areas of the unlabeled and labeled oligonucleotide. (The retention
25 time of fluorescein-labeled dT₁₆ is on the order of 2 to 3 min. longer than the
unlabeled dT₁₆.) The error in this type of assay is about 10%. The percentage
labeling efficiency for 4 types of nucleic acid labeling compounds is shown
below in Table 1.

Example 6Labeled N-(2-hydroxyethoxy)ethyl 2-O-triphosphates (Scheme 3).

Compounds having general formula can be prepared using available reagents and procedures that are standard in the art. In addition, examples of these procedures are illustrated in Schemes 3 and Figure 10.

Scheme 3

10

Example 6ALabeled 2-(2-hydroxyethyl)acetamide 2-O-triphosphates (Figure 10).

[1,4]Dioxan-2-one, (2-Hydroxy-ethoxy)-acetic acid methyl ester, or a polymer thereof, **33**, is reacted with a diamine having a linker moiety, *e.g.*, a polyethylene oxide, alkylene, a combination thereof. The labeled compound is prepared by reacting the ether-amine with a suitable labeling compound using standard conditions. The product is converted into a 5'-triphosphate using standard phosphorylation conditions to afford, respectively, nucleic acid labeling compound **36**.

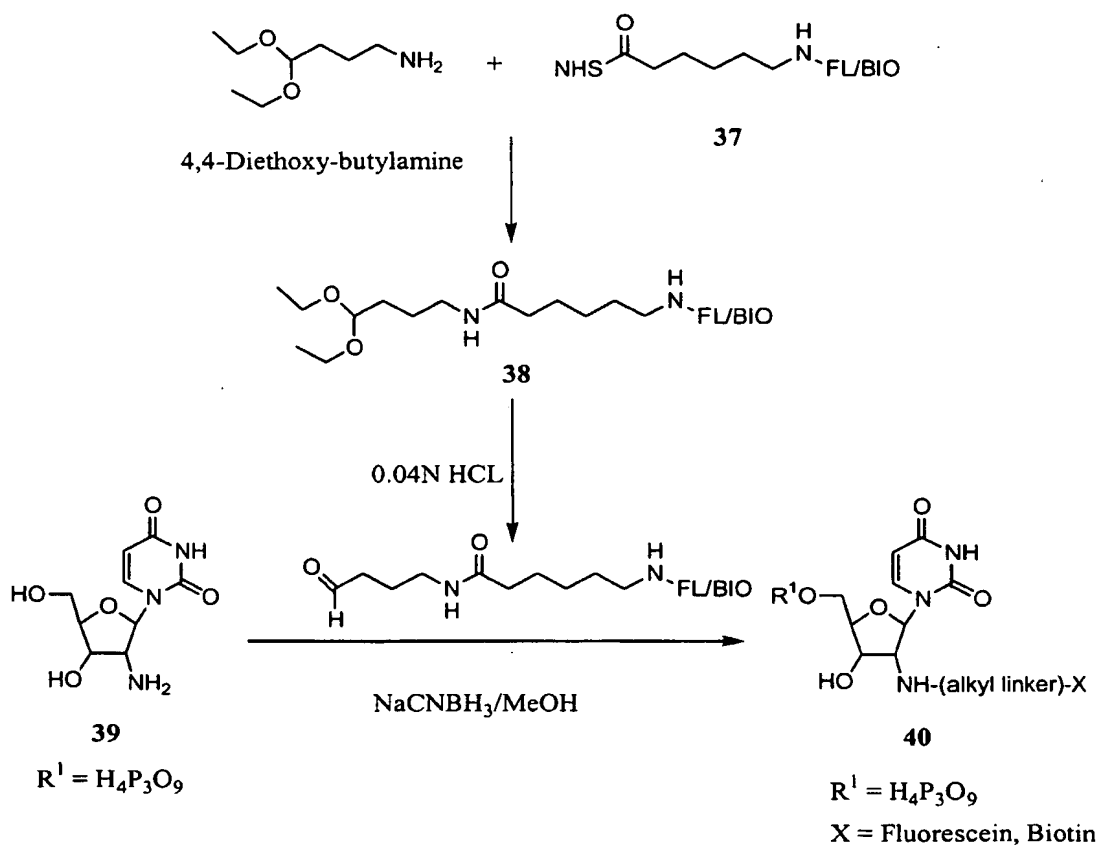
20

Example 7Synthesis of N-alkyl 2'-amino-2'-deoxyuridine triphosphate (Scheme 4).

4,4-Diethoxy-butylamine is reacted with an activated N-labeled caproyl amine, **37** to provide the N-labeled diethoxy butyl caproyl amide, **38**.

- 5 The amide, **38**, is reacted with aqueous hydrochloric acid (0.04 N) to produce the N-labeled aldehyde. the aldehyde is reacted with phosphorylated 1-(3-Amino-4-hydroxy-5-hydroxymethyl-tetrahydro-furan-2-yl)-1H-pyrimidine-2,4-dione, **39**, to furnish labeled N-alkyl 2'-amino-2'-deoxyuridine triphosphate, **40**. The reaction is illustrated in Scheme 4.

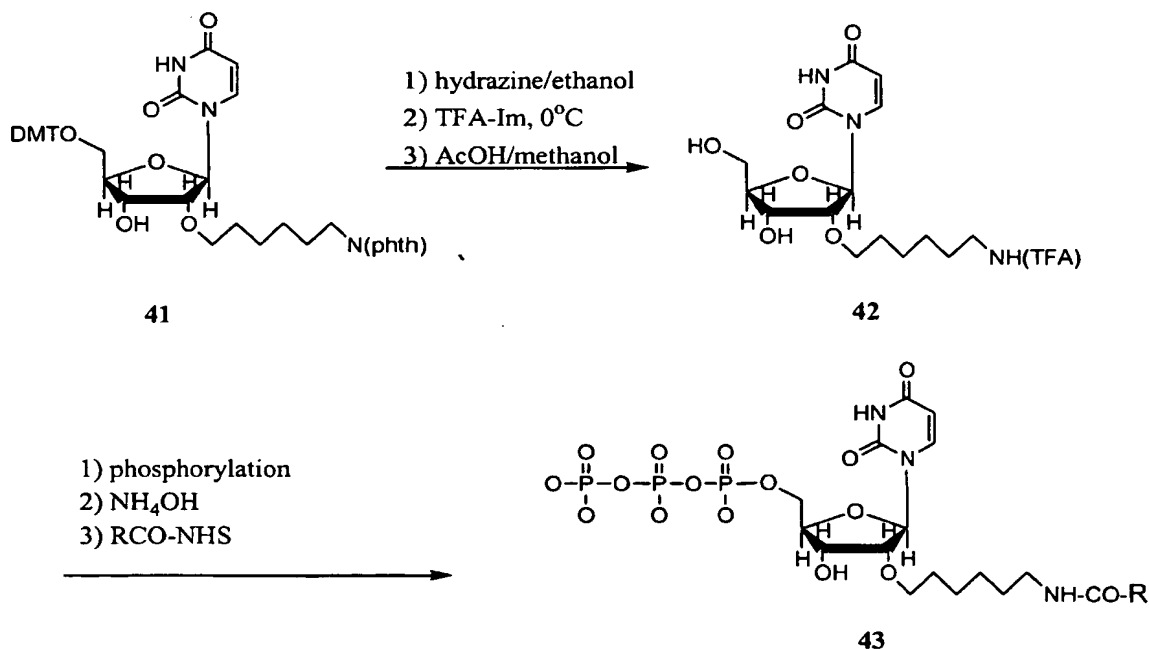
10

Scheme 4

Example 8Synthesis of 2'-O-(6-(Fluorescein-5-carboxamido)hexyl)uridine 5'-O-triphosphate (Scheme 5).

The protected phthalimide, **41**, (available from RI Chemicals, 500 mg), is reacted successively with a) hydrazine/ethanol, b) Trifluoroacetic acid-imidazole at 0°C and c) Acetic acid/methanol to provide compound, **42**. The alcohol-amine is converted into a 5'-triphosphate using standard phosphorylation conditions followed by conjugation to a label to afford, respectively, nucleic acid labeling compounds, **43**.

10

Scheme 5Example 9

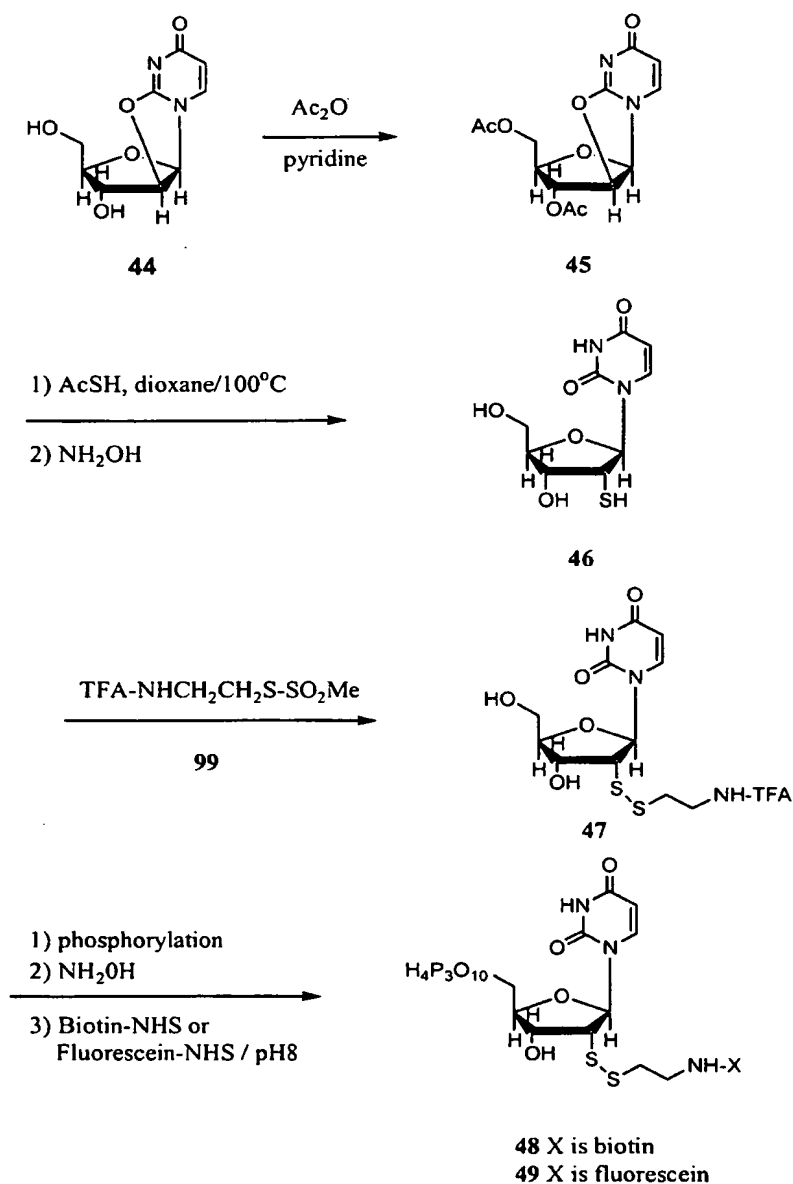
15 Synthesis of 2'-S (N-(6-(Fluorescein-5-carboxamido)hexyl)-aminoethyl)dithiouridine 5'-O- triphosphate (Scheme 6).

3-Hydroxy-2-hydroxymethyl-2,3,3a,9a-tetrahydro-furo[2',3':4,5]oxazolo[3,2-a]pyrimidin-6-one, **44**, is treated with acetic anhydride in pyridine to protect the hydroxy groups. The protected compound, **45**, is reacted with thioacetic acid/dioxane at 100°C, according to the procedure in J.Chem. Soc. Perkin Trans 1, 1997, 2587. the product is treated with ammonia

provide thiol, **46**. The thiol is reacted with disulfide, **99**, to provide disulfide, **47**. the disulfide is converted into a 5'-triphosphate using standard phosphorylation conditions followed by reaction with a labeled reagent to afford, respectively, nucleic acid labeling compounds, **43**.

5

Scheme 6



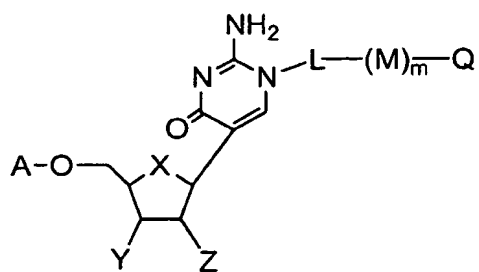
All patents, patent applications, and literature cited in the specification are hereby incorporated by reference in their entirety. In the case of any inconsistencies, the present disclosure, including any definitions therein will prevail.

10

The invention has been described with reference to various specific and preferred embodiments and techniques. However, it should be understood that many variations and modifications may be made while remaining within the spirit and scope of the invention.

Claims

1. A nucleic acid labeling compound having the formula:



wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid;

X is O, S, NR₁ or CHR₂, wherein R₁ and R₂ are, independently, hydrogen, alkyl or aryl;

Y is hydrogen, N₃, F, OR₉, SR₉ or NHR₉, wherein R₉ is hydrogen, alkyl or aryl;

Z is hydrogen, N₃, F, OR₉, SR₉ or NHR₁₀, wherein R₁₀ is hydrogen, alkyl or aryl;

L is a linker moiety; Q is a detectable moiety; and

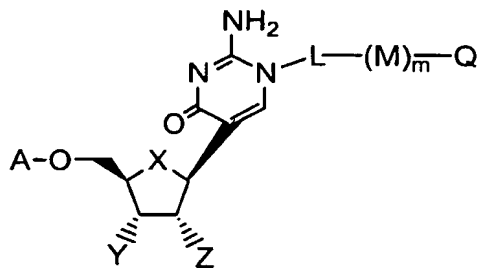
M is a connecting group, wherein m is an integer ranging from 0 to about 20.

2. The nucleic acid labeling compound of claim 1, wherein A is H or H₄O₉P₃⁻; X is O; Y is H or OR₉, wherein R₉ is hydrogen, alkyl or aryl; Z is hydrogen, N₃, F or OR₁₀, wherein R₁₀ is hydrogen, alkyl or aryl; and L is amido alkyl and M is -(M₁)_a-(M₂)_b- wherein a and b are independently integers from 0 to about 5 and the sum of a and b is not greater than 15.
3. The nucleic acid labeling compound of claim 1 or 2, wherein L is -(CH₂)_nC(O)NR₃-, wherein R₃ is hydrogen, alkyl or aryl and n is an integer ranging from about 1 to about 10; M₁ is -(CH₂)_iO- and M₂ is -(CH₂)_jNH-, wherein i and j are independently integers from 1 to about 5.
4. The nucleic acid labeling compound of any of claims 1 to 3, wherein Y is H or OH; Z is H or OH; -L is -CH₂C(O)NH-; M₁ is -(CH₂CH₂O)₃- and M₂ is -CH₂CH₂NH-; and Q is biotin or a carboxyfluorescein.

5. The nucleic acid labeling compound of any of claims 1 to 4, wherein Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{C}(\text{O})\text{NH}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; and Q is biotin.
6. The nucleic acid labeling compound of any of claims 1 to 4, wherein Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{C}(\text{O})\text{NH}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; Q is carboxyfluorescein.
7. The nucleic acid labeling compound of claim 1, wherein A is H or $\text{H}_4\text{O}_9\text{P}_3-$; X is O; Y is H or OR_9 , wherein R_9 is hydrogen, alkyl or aryl; Z is hydrogen, N_3 , F or OR_{10} , wherein R_{10} is hydrogen, alkyl or aryl; and L is amido alkyl and M is $-(M_1)_a-(M_2)_b-(M_3)_c-(M_4)_d-$ wherein a, b, c, and d are independently integers from 0 to about 5 and the sum of a, b, c, and d is not greater than 15.
8. The nucleic acid labeling compound of claim 7, wherein L is $-(\text{CH}_2)_n\text{C}(\text{O})\text{NR}_3-$, wherein R_3 is hydrogen, alkyl or aryl and n is an integer ranging from about 1 to about 10; each M is independently $-\text{C}(\text{O})(\text{CH}_2)_k\text{O}-$, $-(\text{CH}_2)_i\text{O}-$ or $-(\text{CH}_2)_j\text{NH}-$, wherein i, j and k are independently integers from 1 to about 5.
9. The nucleic acid labeling compound of claim 7 or 8, wherein L is $-\text{CH}_2\text{C}(\text{O})\text{NR}_3-$, wherein R_3 is hydrogen or, alkyl; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$ and M_4 is $-(\text{CH}_2)_2\text{NH}-$.
10. The nucleic acid labeling compound of any of claims 7 to 9, wherein Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-\text{C}(\text{O})\text{CH}_2\text{CH}_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; Q is biotin.
11. The nucleic acid labeling compound of any of claims 7 to 9, wherein Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-$; m is 1; and Q is carboxyfluorescein.
12. The nucleic acid labeling compound of claim 7 or 8, wherein L is $-(\text{CH}_2)_2\text{C}(\text{O})\text{NR}_3-$, wherein R_3 is hydrogen or, alkyl; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$ and M_4 is $-(\text{CH}_2)_2\text{NH}-$.
13. The nucleic acid labeling compound of claims 7, 8, or 12, wherein Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-\text{C}(\text{O})\text{CH}_2\text{CH}_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; Q is biotin.
14. The nucleic acid labeling compound of claims 7, 8, or 12, wherein Y is OH; Z is OH; $-L-(M)_m-$ is $-\text{CH}_2\text{CH}_2\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-$; m is 1; and Q is carboxyfluorescein.

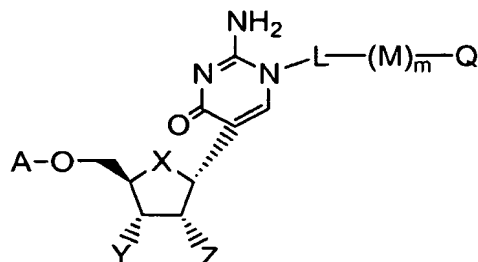
15. The nucleic acid labeling compound of claim 1, wherein A is H or $\text{H}_4\text{O}_9\text{P}_3^-$; X is O; Y is H or OR_9 , wherein R_9 is hydrogen, alkyl or aryl; Z is hydrogen, N_3 , F or OR_{10} , wherein R_{10} is hydrogen, alkyl or aryl; and L is functionalized alkyl and M is $-(\text{M}_1)_a-(\text{M}_2)_b-(\text{M}_3)_c-(\text{M}_4)_d-$ wherein a, b, c, and d are independently integers from 0 to about 5 and the sum of a, b, c, and d is not greater than 15.
16. The nucleic acid labeling compound of claim 15, wherein L is $-(\text{CH}_2)_n\text{O}-$, wherein n is an integer ranging from about 1 to about 10; each M is independently $-\text{C}(\text{O})(\text{CH}_2)_k\text{O}-$, $-(\text{CH}_2)_i\text{O}-$ or $-(\text{CH}_2)_j\text{NH}-$, wherein i, j and k are independently integers from 1 to about 5.
17. The nucleic acid labeling compound of claim 15 or 16, wherein L is $-\text{CH}_2\text{O}-$; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$ and M_4 is $-(\text{CH}_2)_2\text{NH}-$.
18. The nucleic acid labeling compound of claim 15, 16, 17, wherein Y is OH; Z is OH; $-\text{L}-(\text{M})_m-$ is $-\text{CH}_2\text{O}-\text{CH}_2\text{CH}_2\text{NH}-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; Q is biotin.
19. The nucleic acid labeling compound of claim 17, wherein Y is OH; Z is OH; $-\text{L}-(\text{M})_m-$ is $-\text{CH}_2\text{O}-\text{CH}_2\text{CH}_2\text{NH}-$; m is 1; and Q is carboxyfluorescein.
20. The nucleic acid labeling compound of claim 15, wherein L is $-(\text{CH}_2)_n\text{NR}_3-$, wherein R_3 is hydrogen, alkyl or aryl and n is an integer ranging from about 1 to about 10; each M is independently $-\text{C}(\text{O})(\text{CH}_2)_k\text{O}-$, $-(\text{CH}_2)_i\text{O}-$ or $-(\text{CH}_2)_j\text{NH}-$, wherein i, j and k are independently integers from 1 to about 5.
21. The nucleic acid labeling compound of claim 20, wherein L is $-(\text{CH}_2)_6\text{NH}-$; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, and M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$.
22. The nucleic acid labeling compound of claim 21, wherein Y is OH; Z is OH; $-\text{L}-(\text{M})_m-$ is $-(\text{CH}_2)_6\text{NH}-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; and Q is biotin.
23. The nucleic acid labeling compound of claim 21, wherein Y is OH; Z is OH; L is $-(\text{CH}_2)_6\text{NH}-$; m is 0; and Q is carboxyfluorescein.
24. The nucleic acid labeling compound of claim 1, wherein A is H or $\text{H}_4\text{O}_9\text{P}_3^-$; X is O; Y is H or OR_9 , wherein R_9 is hydrogen, alkyl or aryl; Z is hydrogen, N_3 , F or OR_{10} , wherein R_{10} is hydrogen, alkyl or aryl; and L is alkenyl alkyl and M is

- $-(M_1)_a-(M_2)_b-(M_3)_c-(M_4)_d$ wherein a, b, c, and d are independently integers from 0 to about 5 and the sum of a, b, c, and d is not greater than 15.
25. The nucleic acid labeling compound of claim 7, wherein L is - structure $-\text{CH}=\text{CH}-(\text{CH}_2)_n\text{C}(\text{O})\text{N}(\text{R}_6)-$, wherein R_6 is hydrogen, alkyl or aryl and n is an integer ranging from about 0 to about 10; each M is independently $-\text{C}(\text{O})(\text{CH}_2)_k\text{O}-$, $-(\text{CH}_2)_i\text{O}-$ or $-(\text{CH}_2)_j\text{NH}-$, wherein i, j and k are independently integers from 1 to about 5.
 26. The nucleic acid labeling compound of claim 2, wherein L is $-\text{CH}=\text{CH}-\text{C}(\text{O})\text{N}(\text{H})-$; M_1 is $-(\text{CH}_2)_2\text{NH}-$, M_2 is $-\text{C}(\text{O})(\text{CH}_2)_2\text{O}-$, M_3 is $-(\text{CH}_2\text{CH}_2\text{O})_3-$ and M_4 is $-(\text{CH}_2)_2\text{NH}-$.
 27. The nucleic acid labeling compound of claim 12, wherein $-\text{L}(\text{M})_m-$ is $-\text{CH}=\text{CH}-\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-\text{C}(\text{O})\text{CH}_2\text{CH}_2\text{O}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_2\text{CH}_2\text{NH}-$; Y is OH; Z is OH; Q is biotin.
 28. The nucleic acid labeling compound of claim 12, wherein Y is OH; Z is OH; $-\text{L}(\text{M})_m-$ is $-\text{CH}=\text{CH}-\text{C}(\text{O})\text{NH}-\text{CH}_2\text{CH}_2\text{NH}-$; m is 1; and Q is carboxyfluorescein.
 29. The nucleic acid labeling compound of any of claims 1 to 28, having the formula:



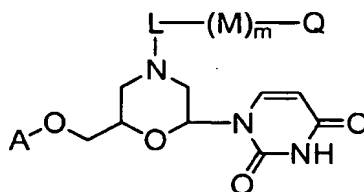
wherein A, X, Y, Z, L, M, m, and Q are as defined herein.

30. A nucleic acid labeling compound of any of claims 1 to 28, having the formula:



wherein A, X, Y, Z, L, M, m, and Q are as defined herein.

31. A nucleic acid labeling compound having the formula:



wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid;

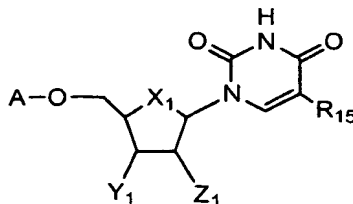
L is a linker moiety;

Q is a detectable moiety; and

M is a connecting group, wherein m is an integer ranging from 0 to about 3.

32. The nucleic acid labeling compound of claim 31, wherein A is H or $\text{H}_4\text{O}_9\text{P}_3^-$; L is alkanoyl; and Q is biotin or a carboxyfluorescein; M is $-\text{NR}_3-$ where R_3 is hydrogen or alkyl, and wherein m is 1 or 0.
33. The nucleic acid labeling compound of claim 32, wherein L is $-\text{C}(\text{O})(\text{CH}_2)_n-$ wherein n is an integer ranging from about 1 to about 10; M is $-\text{NH}-$; Q is biotin or a carboxyfluorescein.
34. The nucleic acid labeling compound of claim 33, wherein $-\text{L}-(\text{M})_m-$ is $-\text{C}(\text{O})(\text{CH}_2)_6-\text{NH}-$; and Q is biotin.

35. The nucleic acid labeling compound of claim 33, wherein $-L-(M)_m-$ is $-C(O)(CH_2)_6-NH-$; and Q is carboxyfluorescein.
36. A nucleic acid labeling compound having the formula:



wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid;

X_1 is O, S, NR_1 or CHR_2 , wherein R_1 and R_2 are, independently, hydrogen, alkyl or aryl;

R_{15} is hydrogen, alkyl or aryl;

Y_1 is hydrogen, N_3 , F, OR_9 , SR_9 or NHR_9 , wherein R_9 is hydrogen, alkyl or aryl;

Z_1 is hydrogen, N_3 , F, OR_9 , SR_9 or NHR_{10} , wherein R_{10} is hydrogen, alkyl or aryl;

wherein one of Y_1 or Z_1 is a group having the formula $-X_2L-(M)_m-Q$;

X_2 is O, S, or NR_{16} , and R_{16} is hydrogen, alkyl or aryl;

L is a linker moiety; Q is a detectable moiety; and

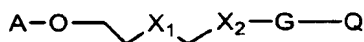
M is a connecting group, wherein m is an integer ranging from 0 to about 20.

37. The nucleic acid labeling compound of claim 36, wherein A is H or $H_4O_9P_3-$; X_1 is O; Y_1 is hydrogen, OR_9 , wherein R_9 is hydrogen, or alkyl, or $-NR_{16}-L-(M)_m-Q$; wherein R_{16} is hydrogen, alkyl or aryl; Z_1 is hydrogen or OR_{10} , wherein R_{10} is hydrogen, or alkyl, or $-NR_{16}-L-(M)_m-Q$, wherein R_{16} is hydrogen, or alkyl; L is N-alkyl amido; R_{15} is hydrogen or alkyl; M is $-(CH_2)_nNR_3-$ where R_3 is hydrogen or alkyl, and m is from 1 to about 15.
38. The nucleic acid labeling compound of claim 37, wherein Y_1 is hydrogen or OH; Z_1 is $-NH-L-(M)_m-Q$; L is $-(CH_2)_nNHC(O)-$ where n is an integer from 1 to about 10; M is $-NH(CH_2)_p-$ where p is an integer from 1 to about 10; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.

39. The nucleic acid labeling compound of claim 38 wherein Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.
40. The nucleic acid labeling compound of claim 39, wherein Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.
41. The nucleic acid labeling compound of claim 39 wherein Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.
42. The nucleic acid labeling compound of claim 39, wherein Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.
43. The nucleic acid labeling compound of claim 37, wherein Y_1 is $-NH-L-(M)_m-Q$; Z_1 is hydrogen or OH; L is $-(CH_2)_nNHC(O)-$ where n is an integer from 1 to about 10; M is $-NH(CH_2)_p-$ where p is an integer from 1 to about 10; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.
44. The nucleic acid labeling compound of claim 43, wherein Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.
45. The nucleic acid labeling compound of claim 43, wherein Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.
46. The nucleic acid labeling compound of claim 43, wherein Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is biotin.
47. The nucleic acid labeling compound of claim 43, wherein Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_4NH-C(O)(CH_2)_5NH-$; and Q is a carboxyfluorescein.
48. The nucleic acid labeling compound of claim 37, wherein A is H or $H_4O_9P_3-$; X_1 is O; Y_1 is hydrogen, OR_9 , wherein R_9 is hydrogen, or alkyl, or $-O-L-(M)_m-Q$; Z_1 is hydrogen OR_{10} , wherein R_{10} is hydrogen, or alkyl, or $-O-L-(M)_m-Q$, or alkyl; L is alkylene; and m is from 1 to about 10.
49. The nucleic acid labeling compound of claim 48, wherein Y_1 is hydrogen or OH; Z_1 is $-O-L-(M)_m-Q$; L is $-(CH_2)_n-$ where n is an integer from 1 to about 12; M is $-NH-$; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.
50. The nucleic acid labeling compound of claim 49, wherein Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.

51. The nucleic acid labeling compound of claim 49, wherein Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is a carboxyfluorescein.
52. The nucleic acid labeling compound of claim 49, wherein Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.
53. The nucleic acid labeling compound of claim 49, wherein Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is a carboxyfluorescein.
54. The nucleic acid labeling compound of claim 37, wherein Y_1 is $-O-L-(M)_m-Q$; Z_1 is hydrogen or OH; L is $-(CH_2)_n-$ where n is an integer from 1 to about 12; M is $-NH-$; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.
55. The nucleic acid labeling compound of claim 54, wherein Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.
56. The nucleic acid labeling compound of claim 54, wherein Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is a carboxyfluorescein.
57. The nucleic acid labeling compound of claim 54, wherein Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NH-$; and Q is biotin.
58. The nucleic acid labeling compound of claim 54, wherein Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-(CH_2)_6-NHC-$; and Q is a carboxyfluorescein.
59. The nucleic acid labeling compound of claim 37, wherein A is H or $H_4O_9P_3-$; X_1 is O; Y_1 is hydrogen, OR_9 , wherein R_9 is hydrogen, or alkyl, or $-S-L-(M)_m-Q$; Z_1 is hydrogen OR_{10} , wherein R_{10} is hydrogen, or alkyl, or $-S-L-(M)_m-Q$; L is alkyl thio; and m is from 1 to about 10.
60. The nucleic acid labeling compound of claim 59, wherein Y_1 is hydrogen or OH; Z_1 is $-S-L-(M)_m-Q$; L is $-S-(CH_2)_n-$ where n is an integer from 1 to about 10; M is $-NH-$; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.
61. The nucleic acid labeling compound of claim 60, wherein Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is biotin.
62. The nucleic acid labeling compound of claim 60, wherein Y_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.

63. The nucleic acid labeling compound of claim 60, wherein Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is biotin.
64. The nucleic acid labeling compound of claim 60, wherein Y_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.
65. The nucleic acid labeling compound of claim 37, wherein Y_1 is $-S-L-(M)_m-Q$; Z_1 is hydrogen or OH; L is $-S-(CH_2)_n-$ where n is an integer from 1 to about 10; M is $-NH-$; R_{15} is hydrogen or methyl; and Q is biotin or a carboxyfluorescein.
66. The nucleic acid labeling compound of claim 65, wherein Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is biotin.
67. The nucleic acid labeling compound of claim 65, wherein Z_1 is OH; R_{15} is hydrogen; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.
68. The nucleic acid labeling compound of claim 65, wherein Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is biotin.
69. The nucleic acid labeling compound of claim 65, wherein Z_1 is OH; R_{15} is methyl; $-L-(M)_m-$ is $-S-(CH_2)_2-NH-$; and Q is a carboxyfluorescein.
70. A nucleic acid labeling compound having the formula:



wherein A is hydrogen or a functional group that permits the attachment of the nucleic acid labeling compound to a nucleic acid;
 X_1 is O, S, NR_1 or CHR_2 , wherein R_1 and R_2 are, independently, hydrogen, alkyl or aryl; X_2 is a bond or alkylene; Q is a detectable moiety; and G is a connecting group.

71. The nucleic acid labeling compound of claim 70, wherein A is H or $H_4O_9P_3-$; X_1 is O; X_2 is a bond; and G is $-C(O)NR_3-(CH_2)_m-(CH_2CH_2O)_n-CH_2CH_2-C(O)NR_3-$, where R_3 is hydrogen or alkyl, and m and n are independently an integer from 1 to about 15.

72. The nucleic acid labeling compound of claim 71, wherein G is $-\text{C}(\text{O})\text{NH}-(\text{CH}_2)_m-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-\text{C}(\text{O})\text{NH}-$ where m is from 1 to about 6 and n is from 1 to about 4.
73. The nucleic acid labeling compound of claim 70, wherein A is H or $\text{H}_4\text{O}_9\text{P}_3^-$; X_1 is O; X_2 is CH_2 ; and G is $-\text{C}(\text{O})\text{NR}_3-(\text{CH}_2)_m-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-\text{C}(\text{O})\text{NR}_3-$, where R_3 is hydrogen or alkyl, and m and n are independently an integer from 1 to about 15.
74. The nucleic acid labeling compound of claim 73, wherein G is $-\text{C}(\text{O})\text{NH}-(\text{CH}_2)_m-(\text{CH}_2\text{CH}_2\text{O})_n-\text{CH}_2\text{CH}_2-\text{C}(\text{O})\text{NH}-$ wherein m is from 1 to about 6 and n is from 1 to about 4.
75. A nucleic acid derivative produced by coupling a nucleic acid labeling compound of claim 2, 7, 12, 15, 20, 24, 32, 37, or 71 with a nucleic acid.
76. A hybridization product, wherein the hybridization product comprises the nucleic acid derivative of claim 75 bound to a complementary probe.
77. The hybridization product of claim 76, wherein the probe is attached to a glass chip.
78. A hybridization product, wherein the hybridization product comprises the nucleic acid derivative of claim 75 bound to a complementary probe.
79. A method of synthesizing a labeled nucleic acid comprising attaching a nucleic acid labeling compound of claim 2, 7, 12, 15, 20, 24, 32, 37, or 71 to a nucleic acid.
80. A method of detecting a nucleic acid comprising incubating a nucleic acid derivative according to claim 75 with a probe.
81. A method according to claim 80, wherein the probe is attached to a glass chip.

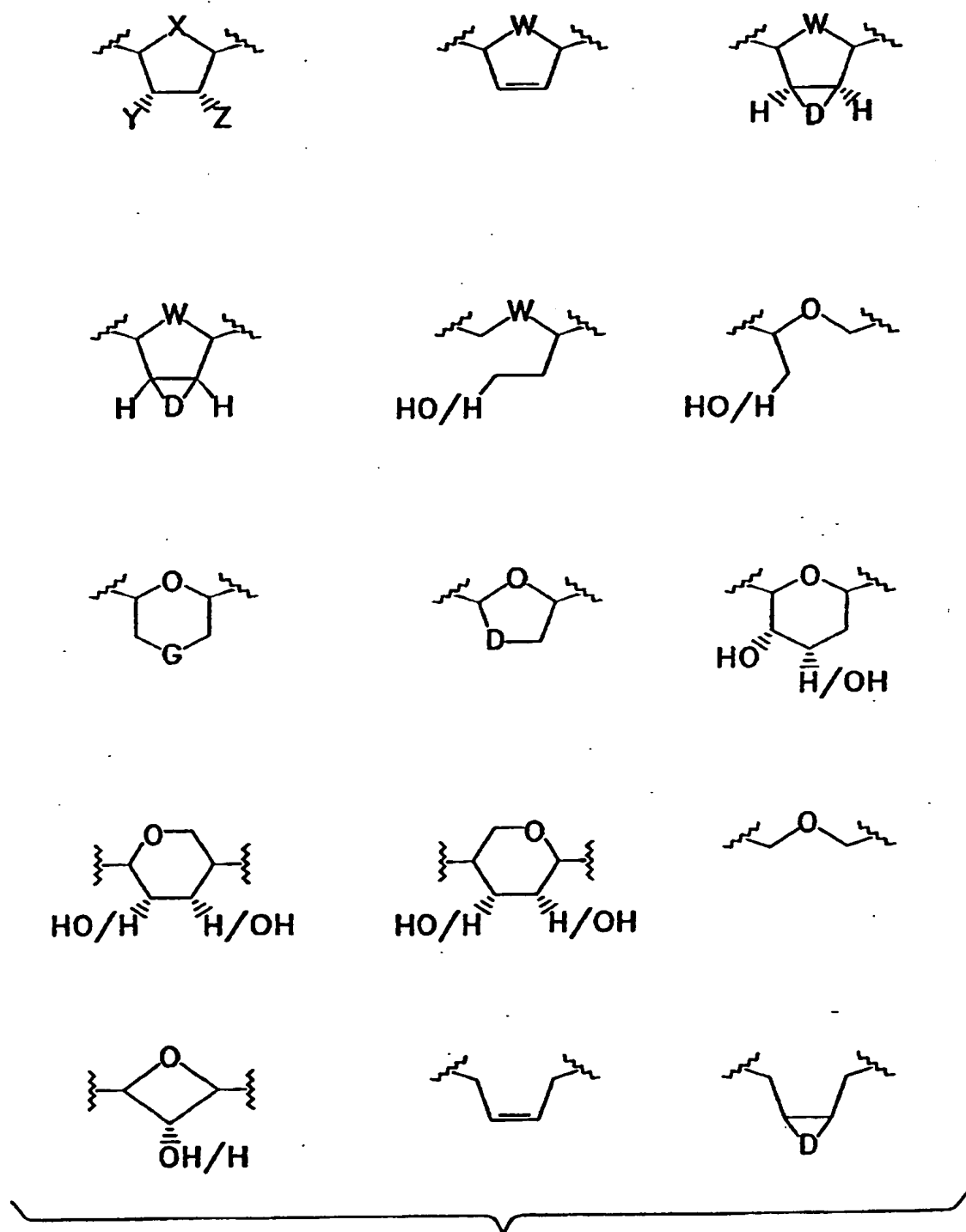
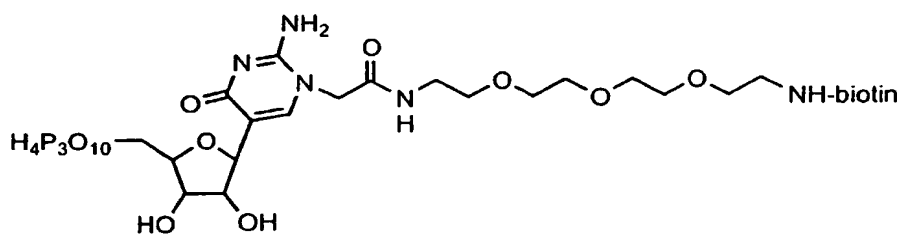
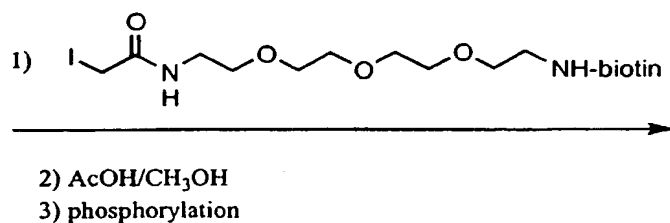
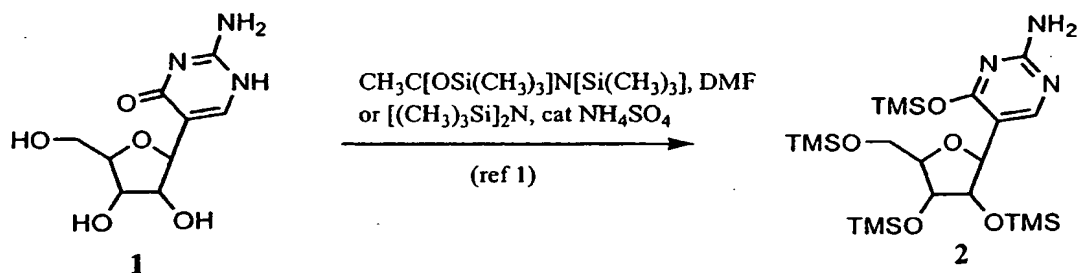


FIG. 1

Fig. 2

fluorescein and biotin labeled N1- ψ CTP

3

Fig. 3

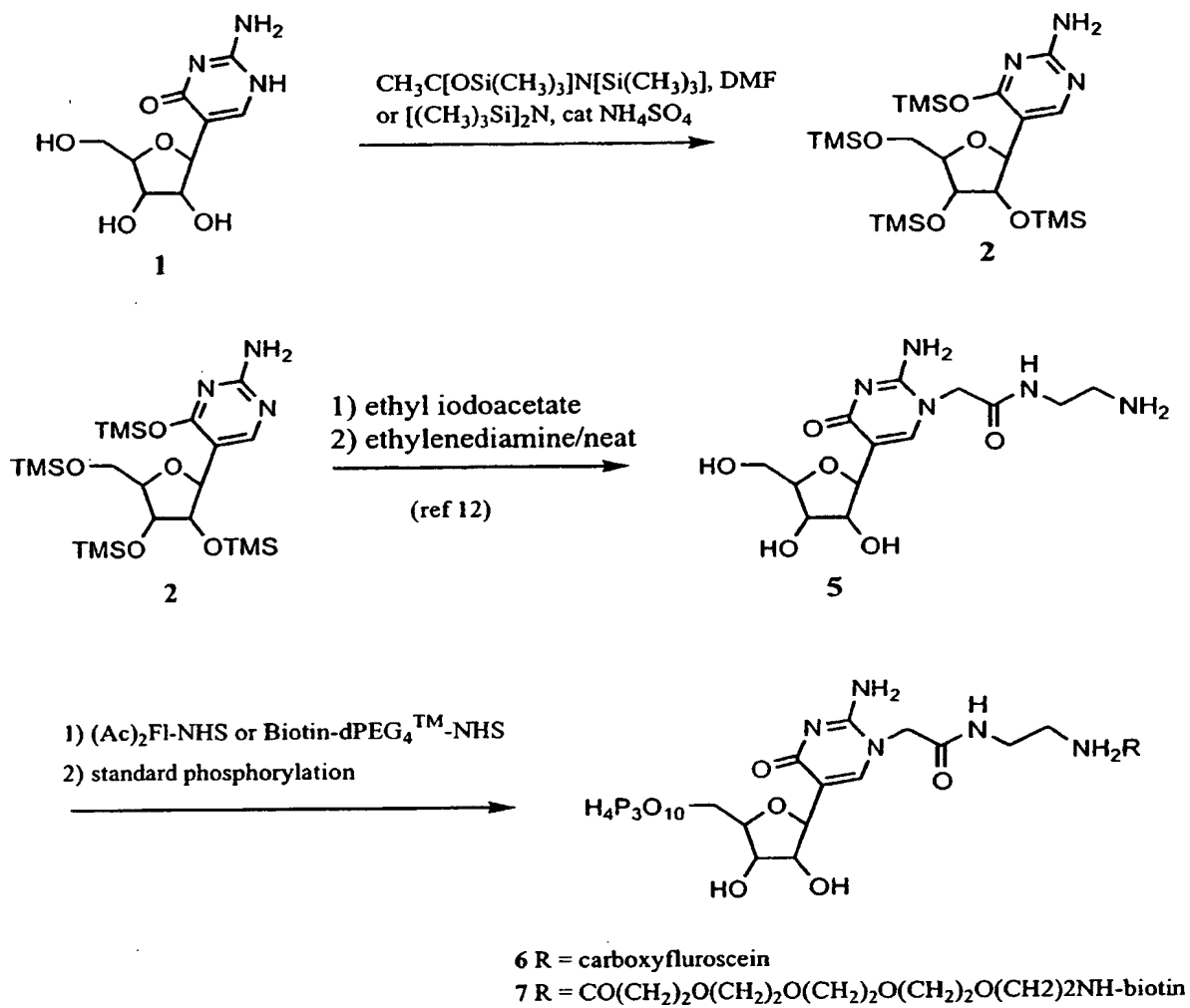


Fig. 3A

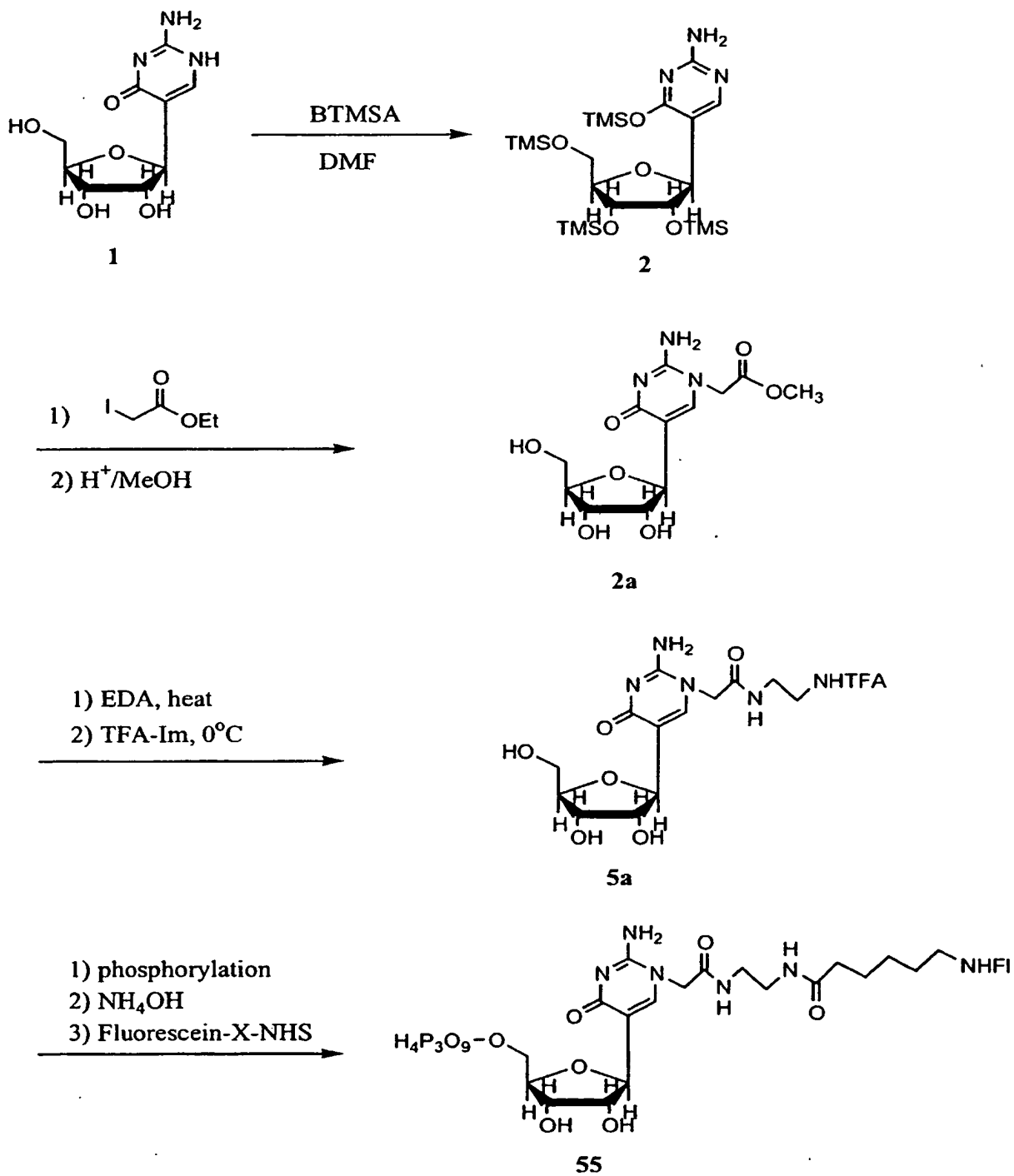


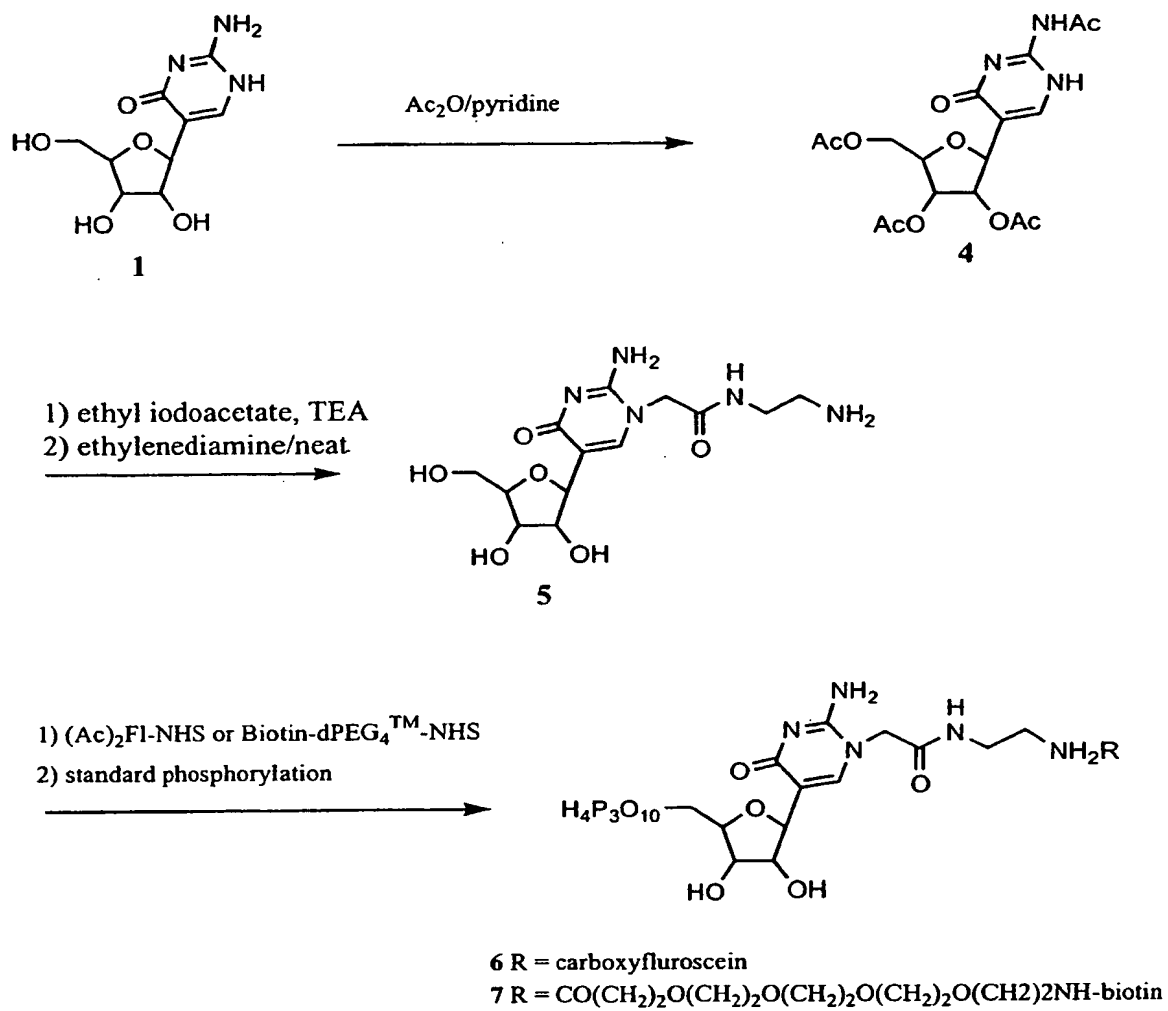
Fig. 4**fluorescein and biotin labeled N1- ψ CTP**

Fig. 5

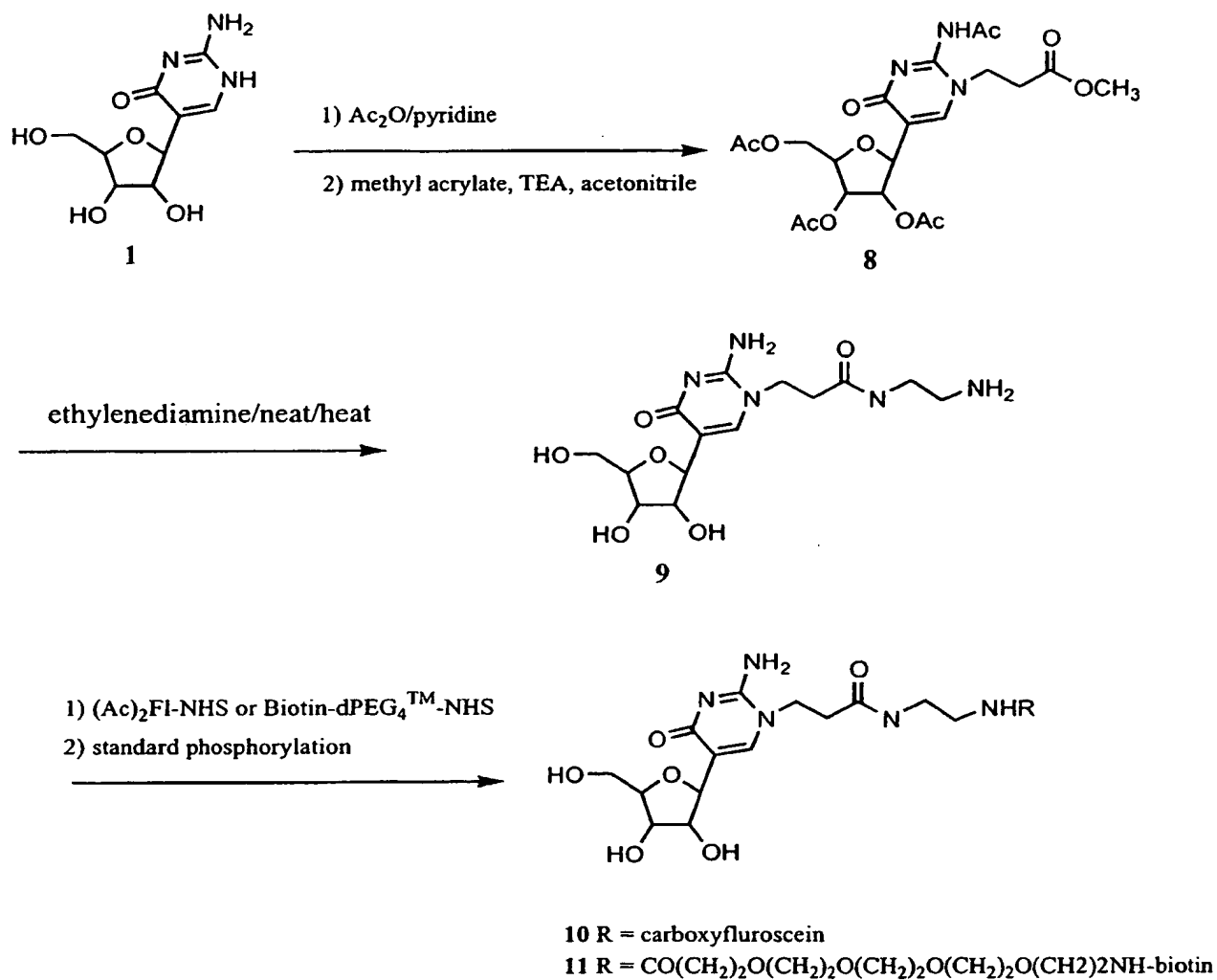
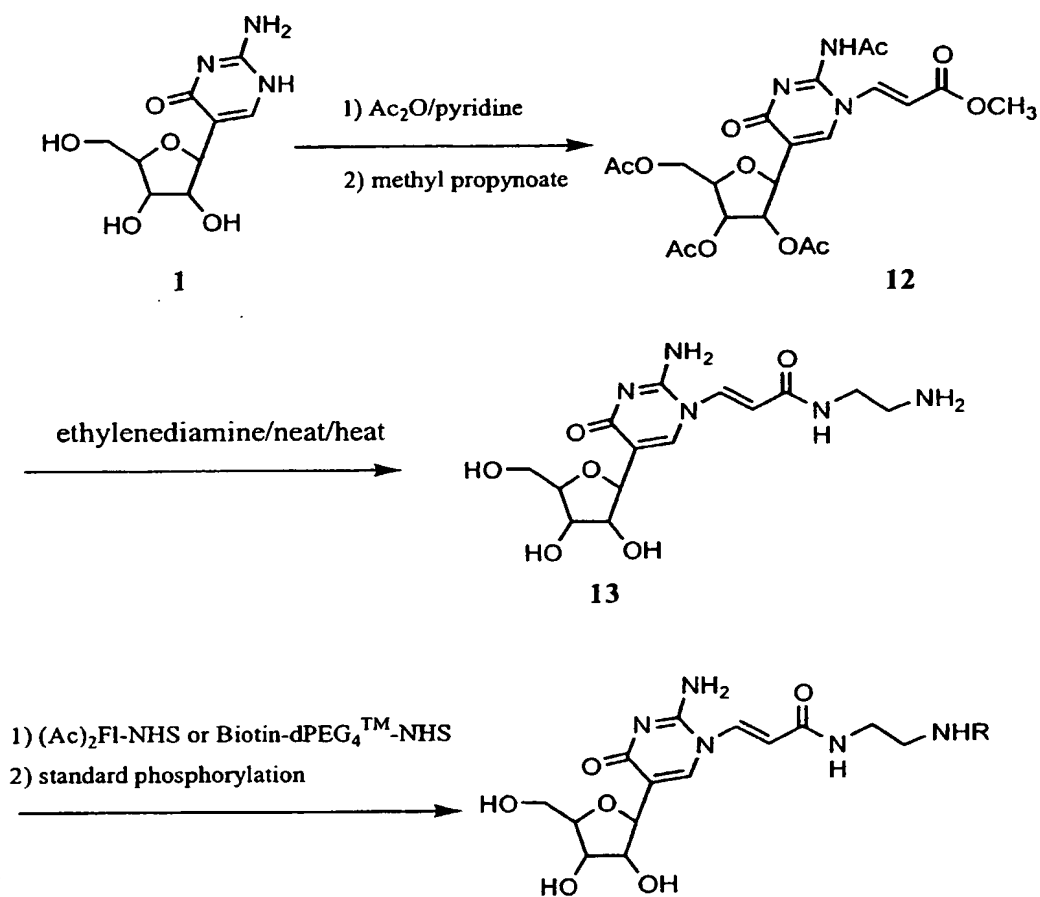


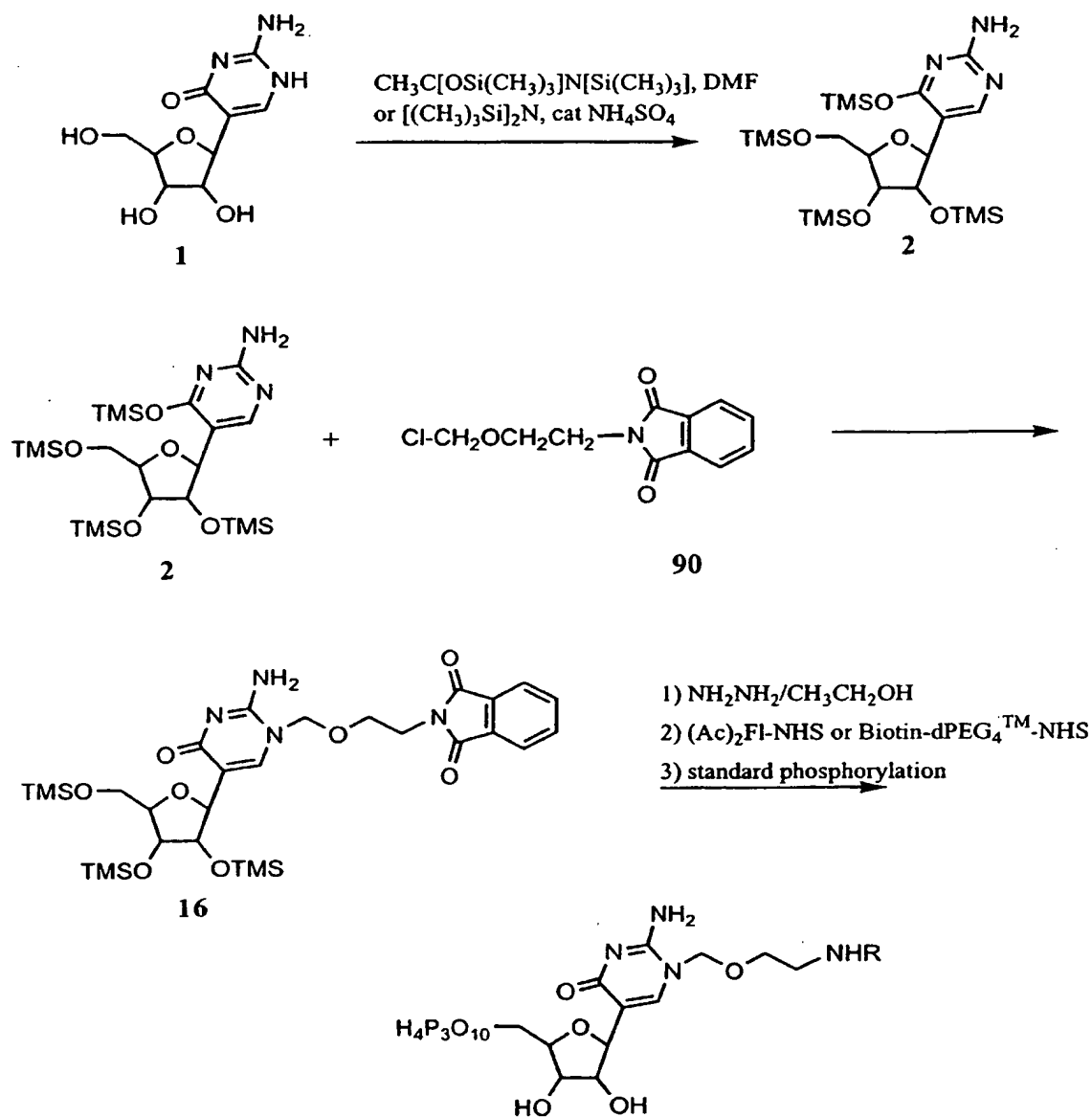
Fig. 6



14 R = carboxyfluorescein

15 R = $\text{CO}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{NH-biotin}$

Fig. 7



17 R = carboxyfluorescein

18 R = $\text{CO}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{O}(\text{CH}_2)_2\text{NH-biotin}$

Fig. 8

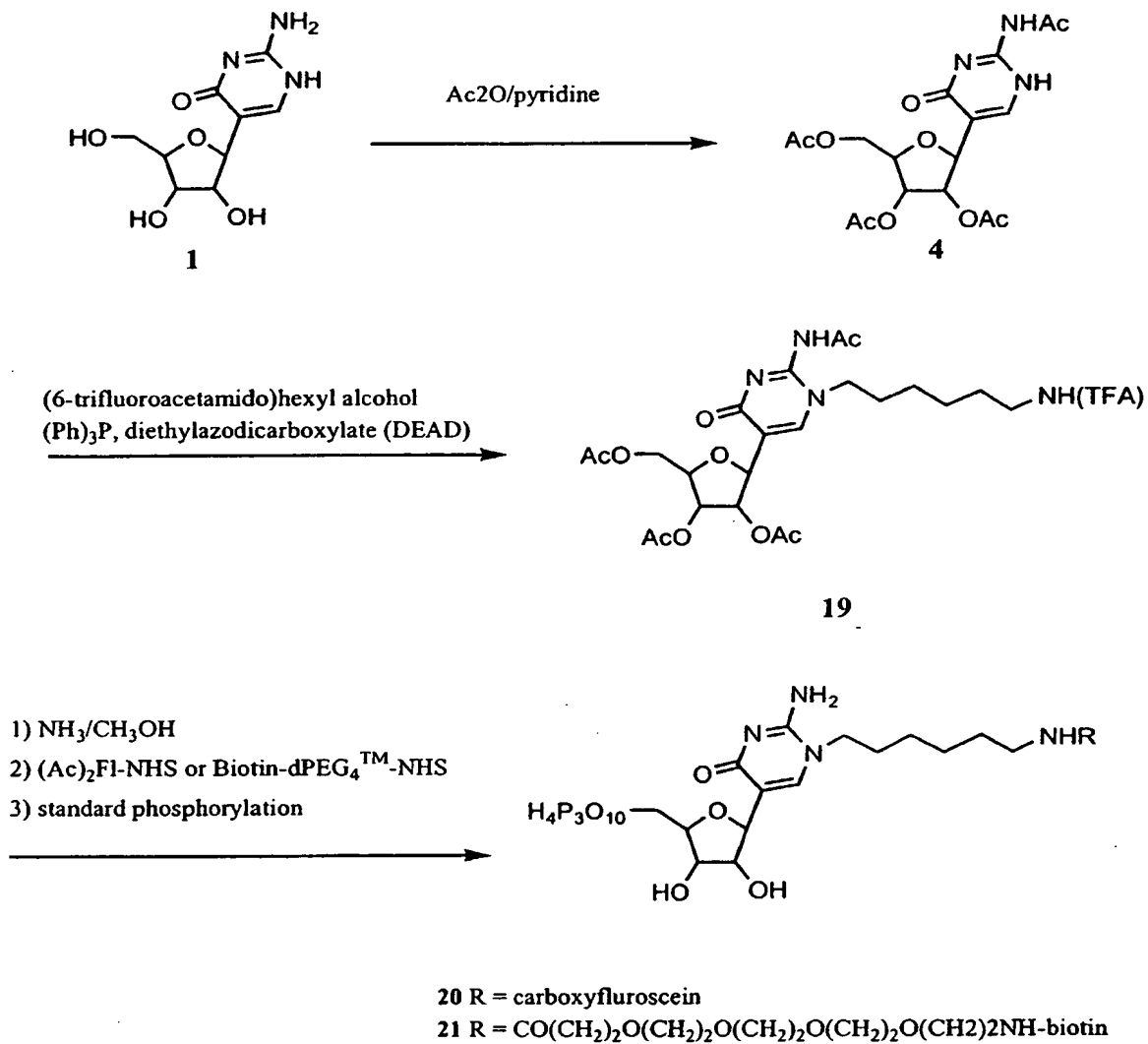


Fig. 9

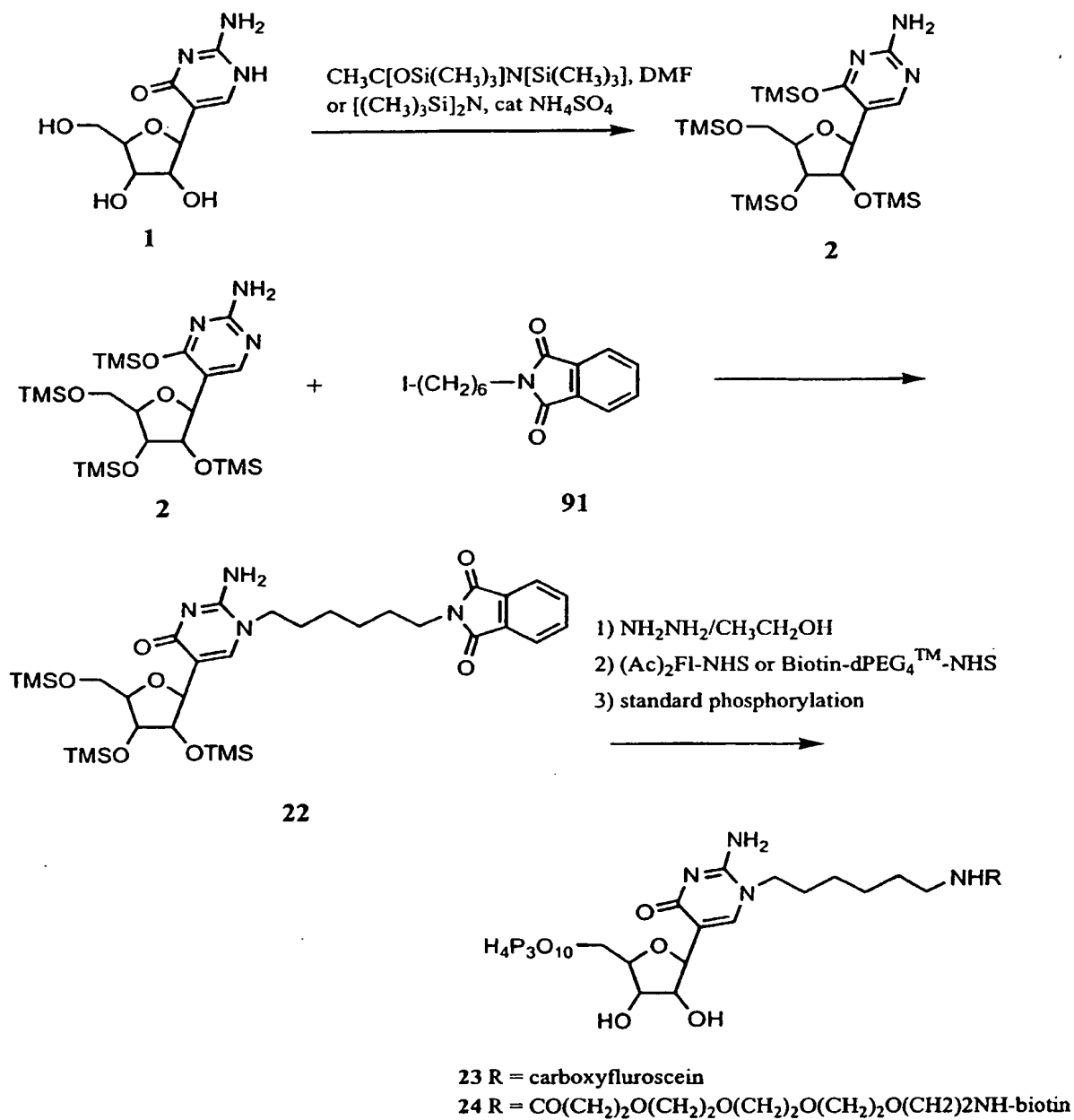


Fig. 10